3 Inner Products, Orthonormal Bases, and Convergence

We'd like to understand how bases work for infinite-dimensional vector spaces. The caveat is that to have a basis for such a space we will need infinitely many basis vectors, and we want to define the span of a set of vectors as the set of linear combinations, something like

$$\operatorname{span}(u_0, u_1, \dots) = \left\{ \sum_{i=0}^{\infty} a_i u_i \,\middle|\, a_i \in \mathbb{R} \text{ for all } i \in \mathbb{Z}_+ \right\} \qquad \text{won't work } \dots$$

This won't work, since we'll need to know when such infinite sums converge, and they probably will not converge for all possible real sequences $a_0, a_1, \dots$.

Finitely non-zero sequences. Define the vector space

$$\ell_{\text{fnz}} = \left\{ x : \mathbb{Z}_+ \to \mathbb{R} \,\middle|\, \text{there exists } N \text{ such that } x_i = 0 \text{ for all } i \geq N \right\}$$

which consists of all discrete-time signals which are zero after some time. Note that signals may have a different numbers of nonzero components.

Now consider the four vector spaces ℓ_e, ℓ_∞, ℓ_2 and ℓ_{fnz}. One might think that a reasonable choice of basis for any of them is

$$\{e_0, e_1, e_2, \dots\}$$

where e_i is the usual vector which has a 1 in the i'th component and zero everywhere else. But this cannot be a basis for all of them; since they are clearly very different spaces. Somehow our notion of basis needs to account for the difference between these spaces. We'll need two things; a notion of convergence and a notion of orthogonality, and will focus on finding a basis for ℓ_2.

3.1 Convergence

Suppose V is a normed vector space. A sequence $x_0, x_1, x_2, \dots$ where each $x_i \in V$ is said to ***converge*** to $a \in V$ if

$$\lim_{k \to \infty} \|x_k - a\| = 0$$

We say the converges with respect to the norm in V. For example, consider the sequence in $L_2([0, \infty), \mathbb{R})$

$$x_k(t) = e^{-kt}$$

converges to 0 in the L_2 norm. Notice that it does not converge with respect to the ∞-norm, and it converges pointwise everywhere except at $t = 0$.

A sequence may converge in $L_2([0, 1])$ without converging pointwise. We'll construct a sequence of functions $x_0, x_1, \dots \in L_2$ as follows. Let m be the integer such that $2^m \leq n < 2^{m+1}$. Divide the interval $[0, 1]$ into 2^m equal sub-intervals, and let $x_n(t) = 1$ for t in the $(n - 2^m + 1)$'th sub-interval and be zero elsewhere. This defines x_n for all $n \in \mathbb{Z}_+$, and $x_n \to 0$ in L_2 but does not converge pointwise.

For the converse example, let y_n be $\sqrt{n}$ on the interval $(0, 1/n)$. Then $\|y_n\|_2 = 1$ for all n but $y_n(t) \to 0$ for all t.

3.2 Open and Closed Sets

Define the ***open unit ball*** around $x \in V$ of radius $\varepsilon > 0$ by

$$B_\varepsilon(x) = \{\, y \in V \mid \|x - y\| < \varepsilon \,\}$$

Suppose S is a subset of V. A point $x \in S$ is called an ***interior point*** of S if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \in S$. A point $x \in V$ is called a ***closure point*** of S if for all $\varepsilon > 0$

$$B_\varepsilon(x) \cap S \neq \emptyset$$

Note that a closure point of S may be outside S.

A set S is called ***open*** if every element of S is an interior point. It is called ***closed*** if every closure point of S is in S. Basic properties are as follows.

(i) S is closed if and only if its complement $\{\, x \in V \mid x \notin S \,\}$ is open.

(ii) If S and T are open, then so is $S \cap T$ and $S \cup T$.

(iii) Suppose Q is a set of subsets of V, and each element of Q is open. Then the union of the elements of Q is open. In other words, a union of infinitely many (even uncountably many) open sets is open. The same is not true for closed sets.

The norm on a vector space determines which sets are open, and determines which sequence converge. In fact once we know which sequences converge we know which sets are open, and vice versa, as follows.

Theorem 1. *A set S is closed if and only if every convergent sequence with elements in S has its limit in S.*

In $\mathbb{C}^n$ every subspace is closed, but in infinite dimensional spaces that is not the case. Consider ℓ_{fnz} as a subspace in ℓ_2. Consider the sequence

$$x_k = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{k}, 0, 0, \ldots\right)$$

where each $x_k \in \ell_{\text{fnz}}$. In the space ℓ_2, the sequence $x_0, x_1, x_2, \ldots$ converges to $a \in \ell_2$. But the limit a is the sequence

$$a = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots,\right)$$

which is not in ℓ_{fnz}. Another way to see this is to pick an element $x \in \ell_2$ which is not in ℓ_{fnz}, and show that x is a closure point of ℓ_{fnz}.

Closed subspaces are useful because we often use iterative algorithms to find minimum solutions to linear equations. If the subspace of solutions to the linear equations is closed, then we can deduce that an algorithm which provides an Cauchy sequence of approximately optimal solutions $x_0, x_1, \ldots$ is actually converging to a solution.

3.3 Completeness

A sequence $x_0, x_1, \ldots$ is called ***Cauchy*** if for all $\varepsilon > 0$ there exists K such that

$$n, m > K \qquad \Longrightarrow \qquad \|x_n - x_m\| \leq \varepsilon$$

This is often written as

$$\lim_{n,m\to\infty} \|x_n - x_m\| = 0$$

A vector space is called ***complete*** if every Cauchy sequence converges. A complete normed space is called a ***Banach Space***. The spaces $\mathbb{C}^n$, ℓ_2 and L_2 are complete.

For subspaces, we can also define completeness; a subspace S is complete if every Cauchy sequence in S converges to a point in S. Then in a Banach space, a subspace is complete if and only if it is closed.

3.4 The Inner Product

Suppose V is a vector space. An ***inner product*** is a function $g : V \times V \to \mathbb{R}$ which satisfies the four conditions below. For two vectors $x, y \in V$, instead of writing $g(x, y)$ we write

$$\langle x, y\rangle = g(x, y)$$

Other notations are used in some books, such as $(x|y)$.

(i) $\langle x, y\rangle = \overline{\langle y, x\rangle}$

(ii) $\langle x, x\rangle = 0$ if and only if $x = 0$

(iii) $\langle x, x\rangle \geq 0$

(iv) $\langle x, \alpha y + \beta z\rangle = \alpha\langle x, y\rangle + \beta\langle x, z\rangle$

Note that on a complex vector space, $\langle x, y\rangle$ is linear with respect to y, but not quite linear with respect to x, since $\langle \alpha x, y\rangle = \overline{\alpha}\langle x, y\rangle$. Some authors use the opposite convention, and make the inner product linear with respect to the first argument.

If $\langle x, y\rangle = 0$ we say x is ***perpendicular*** to y and write $x \perp y$. We define

$$\|x\|^2 = \langle x, x\rangle$$

and this is a norm. A standard property is the Cauchy-Schwartz inequality

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

Notice that here we've defined the norm starting with the inner product. This is not always possible, that is, there are some norms which for which $\|x\|^2 \neq \langle x, x\rangle$ for any inner product. One such norm is the ∞-norm. However, the $2-$norm is constructed from an inner product. On $\mathbb{C}^n$, the inner product is

$$\langle x, y\rangle = x^* y$$

where x^* is the complex-conjugate of the transpose of x. On $L_2([0,1])$ the inner product is

$$\langle x, y\rangle = \int_0^1 \overline{x(t)}y(t)\,dt$$

and on $\ell_2(\mathbb{Z}_+)$ the inner product is

$$\langle x, y\rangle = \sum_{t=0}^{\infty} \overline{x_t}y_t$$

The inner-product is a *continuous* function because

$$4\langle x, y\rangle = \|x+y\|^2 - \|x-y\|^2 - j\|x+jy\|^2 + j\|x-jy\|^2$$

The inner-product provides generalizes the scalar product to arbitrary vector spaces.

A complete inner-product space is called a ***Hilbert Space***.

3.5 The Adjoint

Suppose U and V are Hilbert spaces, each with its own inner-product, and $A : U \to V$ is a bounded linear map. Then there exists a unique linear map $A^* : V \to U$ such that for all $x \in U$ and $y \in V$

$$\langle y, Ax\rangle = \langle A^* y, x\rangle$$

The map A^* is called the ***adjoint*** of A. Existence and uniqueness are straightforward to prove; see, for example, Luenberger or Young. Notice that in the above equation the inner-product on the left-hand side is that in V, and the inner-product on the right-hand side is that in U. The adjoint generalizes the transpose from matrices to more general linear maps. Some key properties are

(i) $\|A\| = \|A^*\|$

(ii) $(AB)^* = B^* A^*$

(iii) $(A^*)^* = A$

(iv) $(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$

A map $T : U \to U$ is called ***Hermitian*** or ***self-adjoint*** if $T^* = T$, and unitary if $T^* T = I$ and $T T^* = I$.