Four

Inner Products and Orthogonal Expansions

In previous section, we used an orthogonality condition to determine the coefficients of a Fourier Sine Series. While at first this may seem to be a clever trick introduced *deus ex machina* to solve a particular problem, it is actually an example of a much deeper idea that motivates the idea of separation of variables. To understand this, we first need to remember some ideas from linear algebra, namely the concepts of an *inner product* and an *orthogonal basis*. We will begin with an example from linear algebra and then generalize the machinery we developed for vectors to deal with functions.

4.1 Orthogonal Bases: An Example in $\mathbb{R}^3$

Consider the following example from linear algebra,

Example 4.1. Solve the linear system

$$\vec{v} = c_1\vec{e}_1 + c_2\vec{e}_2 + c_3\vec{e}_3 = \vec{v} \tag{4.1}$$

where $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$, and $\vec{v}$ are vectors in $\mathbb{R}^3$,

$$\vec{e}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{e}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Solution: The temptation here is rewrite the problem as a linear system

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

and just use gaussian elimination to solve for the unknowns. However the astute reader will notice something about the vectors $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. Remember that we say two vectors $\{\vec{p}, \vec{q}\}$ are perpendicular or *orthogonal* if their dot product is zero, $\vec{p} \cdot \vec{q} = 0$. In fact, the vectors $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ are an *orthogonal basis* because they satisfy the orthogonality condition

$$\vec{e}_i \cdot \vec{e}_j = 0 \qquad \text{if} \quad i \neq j$$

and because they span $\mathbb{R}^3$ (so they form a basis).

We can now solve for each c_i by *projecting* onto the vector $\vec{e}_i$. We can rewrite (4.1) using summation notation,

$$\vec{v} = \sum_{i=1}^{3} c_i \, \vec{e}_i$$

and take the dot product of each side with $\vec{e}_j$ to yield

$$\begin{aligned} \vec{e}_j \cdot \vec{v} &= \vec{e}_j \cdot \left(\sum_{i=1}^{3} c_i \, \vec{e}_i \right) \\ &= \sum_{i=1}^{3} c_i \, \vec{e}_j \cdot \vec{e}_i \\ &= c_j \, \vec{e}_j \cdot \vec{e}_j. \end{aligned}$$

solving for c_j yields

$$\boxed{c_j = \frac{\vec{e}_j \cdot \vec{v}}{\vec{e}_j \cdot \vec{e}_j} = \frac{\vec{e}_j \cdot \vec{v}}{|\vec{e}_j|^2}} \tag{4.2}$$

where we have used the fact that the dot product of a vector with itself is its length squared, $\vec{p} \cdot \vec{p} = |\vec{p}|^2$.

We leave the algebra as an exercise for the reader, but using this formula we discover that

$$c_1 = -\frac{1}{2} \qquad c_2 = 2 \qquad c_3 = -\frac{1}{2}$$

so

$$\boxed{\vec{v} = -\frac{1}{2}\vec{e}_1 + 2\vec{e}_2 - \frac{1}{2}\vec{e}_3.}$$

■

The take-away lesson from this example is that one can expand a function in an orthogonal basis easily by project onto the basis elements. Let us now generalize this idea from vectors to functions.

4.2 Inner Products and Inner Product Spaces

In the previous section we introduced an inner product of two functions u and v on an interval $x \in [a, b]$,

$$\langle u, v \rangle \equiv \int_a^b uv \, dx. \tag{4.3}$$

More generally, a *real inner product* (later we will talk about complex inner products) is a function from two elements of a vector space to the real numbers that satisfies three properties

- **Symmetry:** The inner product is symmetric in its arguments; that is

  $$\langle u, v \rangle = \langle v, u \rangle$$

  for all elements u and v in the vector space.

- **Linearity:** The inner product is linear in each of its arguments; that is

  $$\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle$$

  for all real numbers α and β and any u, v and w in the vector space. Note the symmetry properties implies that linearity in the second argument guarantees linearity in the first argument and vice versa.

- **Positive Definiteness:** We say an inner-product is positive definite if

  $$\langle u, u \rangle \geq 0$$

  with equality when $u = 0$ only[1]. This allows to define the *norm* of a vector,

  $$\|u\| \equiv \sqrt{\langle u, u \rangle}.$$

An *inner product space* is simply a vector space together with an inner product defined on that vector space.

Exercise 4.1. Show the dot product is an inner product for vectors in $\mathbb{R}^n$. Show the norm of a vector with this inner product is its length. Explain why $\mathbb{R}^n$ is an inner-product space.

[1] This assumes the function is continuous; if one chooses to venture into the intricate machinery of measure theory and allow for discontinuous functions one would write "$u = 0$ *almost everywhere*".

Exercise 4.2. Show that

$$\|a-b\|^2 = \|a\|^2 + \|b\|^2 - 2\langle a, b\rangle .$$

Explain how this is related to the law of cosines in $\mathbb{R}^n$.

One very important inner product space is the set of *square-integrable functions*. For our inner product (8.2) above, the set of functions u whose norm is finite,

$$\|u\| \equiv \sqrt{\int_a^b u^2\, dx},$$

is the vector space of square integrable functions, usually refereed to as $L^2[a,b]$. The L here refers to the french mathematician Henri Lebesgue and strictly speaking, the integral should be a *Lebesgue integral* which one generally learns about in a course on measure theory as opposed to the Riemann integral one learns about in calculus. We can safely ignore this nuance for the moment.

4.3 Orthogonal Sets

We can now define an *orthogonal set* of functions,

Definition 4.4. Given a set of non-zero elements $\{\vec{e}_i\}$ of a vector space, $\mathcal{V}$, and an inner product that acts on elements of a vector space, we say the set is an *orthogonal set* if $\vec{e}_i \cdot \vec{e}_j = 0$ if $i \neq j$.

We have already seen two examples of orthogonal sets; in Section (4.1) the vectors

$$\vec{e}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{e}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix},$$

are an orthogonal set in $\mathbb{R}^3$ with respect to the dot product. We also used the orthogonality of the set

$$\{\sin(nx)\} \qquad \text{for} \quad n = 1, 2, 3 \cdots$$

in $L^2[0,\pi]$ to solve for the initial condition of the diffusion heat equation solution we found via separation of variables in Chapter (3) . The amazing thing is that orthogonal sets arise naturally when we apply separation of variables due to the ideas of Sturm-Liouville Theory which we discuss in the next chapter.

Exercise 4.3. Prove the elements of an orthogonal set are linearly independent.

Exercise 4.4. Verify that the following sets are orthogonal

(a) $\{\sin\left(\frac{n\pi x}{\ell}\right)\}$ for $n = 1, 2, 3 \cdots$ in $L^2[0, \ell]$ where $\ell > 0$.

(b) $\{1, x, 3x^2 - 2, 5x^3 - 3x\}$ in $L^2[-1, 1]$.

(c) $\{\cos(nx)\}$ for $n = 0, 1, 2, 3 \cdots$ in $L^2[0, \pi]$.

4.4 Orthogonal Expansions and Least Squares Minimization

A natural question to ask is what is the "best" expansion of a function in terms of a linear combination of an orthogonal set of functions; suppose we are trying to expand a function $f(x)$ in terms of an orthogonal set $\{e_i\}$

$$F \approx \sum_{n=1}^{N} a_n e_n. \tag{4.5}$$

where the a_i are to be determined. A logical choice is to set the projection of both sides of this expression onto the vector e_m equal to each other

$$\begin{aligned}
\langle e_m, F \rangle &\approx \left\langle e_m, \sum_{n=1}^{N} c_n e_n) \right\rangle \\
&\approx \sum_{n=1}^{N} c_n \langle e_m, e_n \rangle \\
&\approx c_m \langle e_m, e_m \rangle \\
&\approx c_m \|e_m\|^2
\end{aligned}$$

which suggests that

$$c_m = \frac{\langle e_m, F \rangle}{\|e_m\|^2}. \tag{4.6}$$

In fact this choice is the best approximation in the *least squares* sense. This is an important result, so we state it as a theorem

Theorem 4.1. *Suppose F is an element of the inner product space $\mathcal{V}$ and $\{e_i\}$ are an orthogonal set in $\mathcal{V}$. The approximation of F by a linear combination of the e_i's,*

$$S_N(x) = \sum_{n=1}^{N} a_n e_n,$$

that minimizes the error

$$E_N = \|F - S_N\|$$

is to choose $a_n = c_n$ *where*

$$c_n \equiv \frac{\langle e_n, F\rangle}{\|e_n\|^2}.$$

Proof. First we note that

$$\begin{aligned}
\|S_N\|^2 &= \langle S_N, S_N\rangle && (4.7)\\
&= \left\langle \sum_{n=1}^{N} a_n e_n, \sum_{m=1}^{N} a_m e_m \right\rangle && (4.8)\\
&= \sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m \langle e_n, e_m\rangle && (4.9)\\
&= \sum_{n=1}^{N} (a_n)^2 \|e_n\|^2. && (4.10)
\end{aligned}$$

Now we expand the square of the error

$$\begin{aligned}
(E_N)^2 &= \|F - S_N\|^2\\
&= \|F\|^2 + \|S_N\|^2 - 2\langle F, S_N\rangle\\
&= \|F\|^2 + \sum_{n=1}^{N}(a_n)^2\|e_n\|^2 - 2\left\langle F, \sum_{n=1}^{N} a_n e_n\right\rangle \qquad \boxed{\text{Using (4.10)}}\\
&= \|F\|^2 + \sum_{n=1}^{N}(a_n)^2\|e_n\|^2 - 2\sum_{n=1}^{N} a_n \langle F, e_n\rangle\\
&= \|F\|^2 + \sum_{n=1}^{N}(a_n)^2\|e_n\|^2 - 2\sum_{n=1}^{N} a_n c_n \|e_n\|^2\\
&= \|F\|^2 + \sum_{n=1}^{N}[(a_n)^2 - 2a_n c_n]\|e_n\|^2\\
&= \|F\|^2 + \sum_{n=1}^{N}\left[(a_n - c_n)^2 - c_n^2\right]\|e_n\|^2
\end{aligned}$$

This expression is minimized when the terms proportional to $(a_n - c_n)^2$ vanish, that is when we choose $a_n = c_n$, which completes the proof. □

A nice corollary to this result is that with this choice of coefficients, the square of the error can be written as

$$(E_N)^2 = \|F\|^2 - \sum_{n=1}^{N}(c_n)^2\|e_n\|^2. \qquad (4.11)$$

4.5 Bessel's Inequality and Parseval's Identity

Because the square of the error is positive, $(E_N)^2$ an immediate consequence of (4.12) is *Bessel's Inequality*,

$$\|F\|^2 \geq \sum_{n=1}^{N} (c_n)^2 \|e_n\|^2. \tag{4.12}$$

Bessel's inequality can be thought of as a measure of the goodness of the least squares approximation. We can now define when an *orthogonal set* is an *orthogonal basis*. The term basis employs that the set is complete; that is any element in the vector space can be written as a linear combination of the elements of the set. For a finite dimensional space, such as $\mathbb{R}^n$, this definition is clear. However, the space $L^2[a,b]$ is infinite dimensional, so one must be a little more careful. We say that an orthogonal set, $\{e_i\}$, is complete in L^2 if

$$\lim_{N\to\infty} E_N = 0.$$

In this case Bessel's Inequality becomes *Parseval's Identity*,

$$\|F\|^2 = \sum_{n=1}^{\infty} (c_n)^2 \|e_n\|^2. \tag{4.13}$$

which is true for any orthogonal basis in L^2. This can be a useful practical check of completeness and convergence of an expansion, as we will see in our discussion of Fourier series.

Exercise 4.5. In $\mathbb{R}^N$, Parseval's Identity states that

$$\|F\|^2 = \sum_{n=1}^{N} (c_n)^2 \|e_n\|^2.$$

Verify this identity for the $\mathbb{R}^3$ example in Section (4.1). Explain how this result relates to Pythagoras' Theorem in $\mathbb{R}^N$.