

Orthonormal Functions

Objective: To understand the notion of orthogonal and orthonormal functions, and how to use these notions to decompose a function into a sum of orthonormal functions.

Warm-up: Orthonormal Decomposition:

Consider the interval $[0, 1]$ on the real x -axis. Let $f(x) = 1$, $g(x) = \sqrt{3}(2x - 1)$. For any two integrable functions $\phi(x)$ and $\psi(x)$ on this interval, let

$$\langle \phi | \psi \rangle = \int_0^1 \phi(x)\psi(x) \, dx.$$

Note on pronunciation: the left-hand side of the above equation can be read (amongst other possibilities) as “phi bracket psi”, or “bra-phi, ket-psi.”

1. Show that for any three integrable functions ϕ , ψ , and η on the interval $[0, 1]$, and constants a and b :

$$\langle a\phi + b\psi | \eta \rangle = a\langle \phi | \eta \rangle + b\langle \psi | \eta \rangle,$$

and

$$\langle \phi | \psi \rangle = \langle \psi | \phi \rangle.$$

2. Show that

$$\langle f | g \rangle = 0.$$

3. Show that

$$\langle f | f \rangle = \langle g | g \rangle = 1.$$

4. Suppose $3x - 5$ can be written as $\alpha f(x) + \beta g(x)$, for some constants α and β . We will find these constants.

(a) Compute $\langle 3x - 5 | f \rangle$.

(b) Compute $\langle 3x - 5 | g \rangle$.

(c) Show that $3x - 5 = \langle 3x - 5 | f \rangle f(x) + \langle 3x - 5 | g \rangle g(x)$.

Theory: Inner Products and Orthogonality:

An **inner product** of functions is a rule that takes two functions ϕ and ψ and returns a real number, often denoted $\langle \phi | \psi \rangle$, and satisfies the following three conditions for any functions ϕ , ψ , and η and constants a and b :

$$\begin{aligned}\langle a\phi + b\psi | \eta \rangle &= a\langle \phi | \eta \rangle + b\langle \psi | \eta \rangle && \text{linearity in 1st argument} \\ \langle \phi | \psi \rangle &= \langle \psi | \phi \rangle && \text{symmetry} \\ \langle \phi | \phi \rangle &> 0 && \text{positive-definiteness.}\end{aligned}$$

For the last condition, we need $\phi \neq 0$; it follows from linearity that $\langle 0 | 0 \rangle = 0$.

Inner products are a more general notion than we present here; if you've ever seen the “dot product” of vectors, that is an example of an inner product. However, we shall only be interested in inner products of functions for the purposes of Fourier analysis. There are many different inner products one can define, but we will be particularly focused on ones arising from integrals. If we have a class of integrable functions on an interval $[a, b]$, then defining

$$\langle \phi | \psi \rangle = \int_a^b \phi(x)\psi(x) \, dx$$

gives an inner product (up to a slight issue with identifying functions that agree “almost everywhere.”

Given an inner product, the **norm** of a function ϕ (with respect to that inner product), denoted $\|\phi\|$, is defined to be

$$\|\phi\| = \sqrt{\langle \phi | \phi \rangle}.$$

Because of positive-definiteness, this is guaranteed to be a real number, and is strictly positive except when $\phi = 0$. Note that the root-mean-square of a function is precisely the norm when the inner product used is an integral, as in our case. In general, the norm can be regarded as a measure of the “size” of a function.

If ϕ_1, \dots, ϕ_n are a collection of functions, and $\langle \phi_i | \phi_j \rangle = 0$ whenever $i \neq j$, we say that these functions are **orthogonal**. If also $\|\phi_i\| = 1$ for all i , then we say they are **orthonormal**. A convenient notation for this is the **Kronecker delta**, defined by:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Then ϕ_1, \dots, ϕ_n are orthonormal if $\langle \phi_i | \phi_j \rangle = \delta_{ij}$.

We can now see that the functions f and g of the warm-up were orthonormal with respect to the integral inner product on the interval $[0, 1]$.

Theory: Linear Combinations and Measuring “Distances”:

In the warm-up we had two orthonormal functions f and g ; we then took a third function, $3x - 5$, and wrote it in the form $\alpha f(x) + \beta g(x)$, where α and β were given by $\alpha = \langle f \mid 3x - 5 \rangle$, $\beta = \langle g \mid 3x - 5 \rangle$. This is a general phenomenon, and works for any number of orthonormal functions. If ϕ_1, \dots, ϕ_n are functions (not necessarily orthonormal), we say that another function ψ is a **linear combination** of ϕ_1, \dots, ϕ_n if there are constants a_1, \dots, a_n such that

$$\psi(x) = \sum_{i=1}^n a_i \phi_i(x).$$

For example, a degree- n polynomial is a linear combination of the functions $1, x, \dots, x^n$, and the product-to-sum formulae allow us to express products of sinusoids as linear combinations of sinusoids.

If we have an inner product and the functions ϕ_1, \dots, ϕ_n are orthonormal, then we can easily find the values of the coefficients a_i in any linear combination of the ϕ_i :

Show that if

$$\psi(x) = \sum_{i=1}^n a_i \phi_i(x),$$

then $a_i = \langle \psi \mid \phi_i \rangle$.

Another use of inner products is that they give us a notion of “distance between functions.” This comes from the idea of viewing the norm as a measure of the “size” of a function: the distance between two functions ϕ and ψ should be the size of the difference $\psi(x) - \phi(x)$; so we define the distance from ψ to ϕ to be the real number $\|\phi - \psi\|$. Note that, by positive-definiteness, $\|\phi - \psi\|$ is strictly positive, unless $\phi = \psi$, in which case it is 0. This agrees with our intuition about distance.

If the inner product we are using is the integral inner product over some interval, then $\|\phi - \psi\|$ is the root-mean-square of $\phi(x) - \psi(x)$, so is essentially the average difference between $\phi(x)$ and $\psi(x)$.

Application: Orthonormal Approximations:

We have seen that if ϕ_1, \dots, ϕ_n are orthonormal with respect to some inner product, then given any function ψ which can be written as a linear combination of the ϕ_i , we can use the inner product to find the coefficients of this linear decomposition: they are $\langle \psi | \phi_i \rangle$. However, this assumes that ψ can be written as a linear combination of the ϕ_i , which means it wasn't that different a function from them to begin with.

We shall show that even if ψ *cannot* be written as a linear combination of the ϕ_i , we can use this same calculation to find the best possible approximation to ψ by a linear combination of the ϕ_i . This means that if we have an orthonormal collection of functions ϕ_i which we understand well and a function ψ which we want to understand, we can find the best approximation to ψ by the ϕ_i . We shall see that the key idea behind Fourier analysis is to approximate functions with sinusoids; since sinusoids are well understood, if a complicated function can be approximated by sinusoids, we can hope to understand it through this approximation.

What do we mean though by “best possible” approximation? We mean best possible *with respect to the chosen inner product*. That is, we find a function $\Psi(x)$ which is a linear combination of the ϕ_i and is such that $\|\Psi - \psi\|$ is as small as possible. A different inner product would give us a different notion of orthogonality, and a different measure of the closeness of an approximation.

We use a mathematical technique called **induction** which is useful for proving statements involving sequences (like our ϕ_1, \dots, ϕ_n). Induction is a step-by-step process: first we find the best approximation to ψ by a multiple of ϕ_1 , then show how to modify this to find the best approximation by a linear combination of ϕ_1 and ϕ_2 , and so on. So we will first prove that $\langle \phi_1 | \psi \rangle \phi_1$ is the closest approximation to ψ by a multiple of ϕ_1 , and then that if Ψ_k is the closest approximation to ψ by a linear combination of ϕ_1, \dots, ϕ_k , then $\Psi_k + \langle \phi_{k+1} | \psi \rangle \phi_{k+1}$ is the closest approximation to ψ by a linear combination of $\phi_1, \dots, \phi_{k+1}$. We can then take this to be our Ψ_{k+1} , and continue.

Note that we work purely in terms of an inner product. Really, this has nothing to do with integrals, or even functions, but is purely about using an inner product to find a best approximation to a given object by a linear combination of orthonormal objects.

Application: Orthonormal Approximations (cont.):

Let ϕ_1, \dots, ϕ_n be orthonormal and ψ be a function we wish to approximate. We prove that the linear combination

$$\Psi = \sum_{i=1}^n \lambda_i \phi_i$$

which minimises $\|\Psi - \psi\|$ (*i.e.* gives the best approximation) is the one where $\lambda_i = \langle \psi | \phi_i \rangle$ for each i .

1. First we show that

$$\left\| \psi - \sum_{i=1}^n \lambda_i \phi_i \right\|^2 = \|\psi\|^2 - 2 \sum_{i=1}^n (\lambda_i \langle \psi | \phi_i \rangle) + \sum_{i=1}^n \lambda_i^2. \quad (\star)$$

This may seem intimidating, but we'll break it down and get there.

(a) First prove that for any function η and any k :

$$\|\eta - \lambda_k \phi_k\|^2 = \|\eta\|^2 - 2\lambda_k \langle \eta | \phi_k \rangle + \lambda_k^2.$$

(b) Take $\eta = \psi$ and $k = 1$ to prove Equation (\star) in the case $n = 1$.

(c) Now take

$$\eta = \psi - \sum_{i=1}^{k-1} \lambda_i \phi_i$$

and show that, for this η , $\langle \eta | \phi_k \rangle = \langle \psi | \phi_k \rangle$. Hence conclude that if Equation (\star) is true for $n = k - 1$, then it is true for $n = k$.

(d) Now we know that (\star) is true for $n = 1$, and if true for $n = k - 1$, is also true for $n = k$. Why does this mean it is true for all n ?

2. By Equation (\star) , in order to minimise

$$\left\| \psi - \sum_{i=1}^n \lambda_i \phi_i \right\|^2,$$

it suffices to minimise $\lambda_i^2 - 2\lambda_i \langle \psi | \phi_i \rangle$ for each value of i . Why does minimising the square of the norm also minimise the norm itself?

3. Use Fermat's Method and the 2nd derivative test to show that $\lambda_i^2 - 2\lambda_i \langle \psi | \phi_i \rangle$ has its minimum value when $\lambda_i = \langle \psi | \phi_i \rangle$. This completes the proof.

Practice:

1. Prove (from the three defining conditions) that for any inner product, we have linearity in the 2nd argument: $\langle \phi \mid a\psi + b\eta \rangle = a\langle \phi \mid \psi \rangle + b\langle \phi \mid \eta \rangle$.
2. Check the three conditions to prove that the following are inner products of functions on the interval $[a, b]$:

(a)

$$\langle \phi \mid \psi \rangle = \phi(a)\psi(a).$$

(b)

$$\langle \phi \mid \psi \rangle = \int_a^b \phi(x)\psi(x) \, dx.$$

3. Consider points in \mathbb{R}^3 (real, 3-dimensional space, so points have (x, y, z) -coordinates); define a pairing by

$$\langle (x_1, y_1, z_1) \mid (x_2, y_2, z_2) \rangle = x_1x_2 + y_1y_2 + z_1z_2.$$

This is called the **dot product**, and often written $(x_1, y_1, z_1) \cdot (x_2, y_2, z_2)$.

- (a) Show that the dot product is an inner product, by checking the three conditions. Note: for points in \mathbb{R}^3 , addition and multiplying by a constant are done coordinate-wise; so $(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$ and $a(x, y, z) = (ax, ay, az)$.

- (b) Show that

$$\phi_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right), \quad \phi_2 = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0 \right)$$

are orthonormal with respect to the dot product.

- (c) Let ψ be the point $(-1, 7, 4)$. Find the best approximation to ψ by a linear combination of ϕ_1 and ϕ_2 , and the error in this approximation (*i.e.*, the distance from this approximation to ψ).
4. Consider integrable functions from 0 to 2π , with the inner product given by integration (as in part (c) of question 1).
 - (a) Show that $\phi_1 = \frac{1}{\sqrt{\pi}} \cos(t)$ and $\phi_2 = \frac{1}{\sqrt{\pi}} \cos(3t)$ are orthonormal.
 - (b) Let $\psi(t) = \sin(t) \sin(2t)$. Given that $\langle \psi \mid \phi_1 \rangle = \frac{\sqrt{\pi}}{2}$ and $\langle \psi \mid \phi_2 \rangle = -\frac{\sqrt{\pi}}{2}$, write down the closest approximation to ψ by a linear combination of ϕ_1 and ϕ_2 .
 - (c) Apply a product-to-sum formula to ψ to verify your result from part (b).

Key Points to Remember:

1. An **inner product** is a pairing which takes two inputs (functions, points in space, etc.) and gives a real number output, satisfying the following three conditions:

$$\begin{aligned}\langle a\phi + b\psi \mid \eta \rangle &= a\langle \phi \mid \eta \rangle + b\langle \psi \mid \eta \rangle && \text{linearity in 1}^{\text{st}} \text{ argument} \\ \langle \phi \mid \psi \rangle &= \langle \psi \mid \phi \rangle && \text{symmetry} \\ \langle \phi \mid \phi \rangle &> 0 && \text{positive-definiteness.}\end{aligned}$$

2. Given an inner product, the **norm** of an object ϕ is $\|\phi\| = \sqrt{\langle \phi \mid \phi \rangle}$. This is 0 if $\phi = 0$, and otherwise is strictly positive.
3. The **distance** between two objects ϕ and ψ is the size of their difference, $\|\phi - \psi\|$.
4. A **linear combination** of objects ϕ_1, \dots, ϕ_n is any expression of the form

$$\sum_{i=1}^n a_i \phi_i,$$

where the a_i are constants.

5. We say objects ϕ_1, \dots, ϕ_n are **orthogonal** if $\langle \phi_i \mid \phi_j \rangle = 0$ for all $i \neq j$. If also $\|\phi_i\| = 1$ for all i , we say they are **orthonormal**. We can concisely combine the two conditions for orthonormality by writing $\langle \phi_i \mid \phi_j \rangle = \delta_{ij}$ (*i.e.*, is 1 if $i = j$ and 0 if $i \neq j$).
6. Given ϕ_1, \dots, ϕ_n orthonormal, and another object ψ , we can find the best approximation to ψ by a linear combination of the ϕ_i ; it is

$$\sum_{i=1}^n \langle \psi \mid \phi_i \rangle \phi_i.$$