Big vector spaces

September 26, 2019

1 Vector spaces

You are familiar with the idea of a vector: an object much like an arrow that has *magnitude* and *direction*. These intuitive descriptors are indeed crucial properties of vectores, but we need a more formal understanding of them.

Here we identify some basic formal properties of vector spaces. We will use these to describe solutions to differential equations in greater detail.

1.1 Vector space: a formal definition

Briefly, a *vector space* is a set of objects, $\mathcal{S}$, which we may (1) add and (2) take scalar multiples.

1. By addition, denoted as usual
$$\vec{\mathbf{u}} + \vec{\mathbf{v}}$$
we mean a way of combining any two vectors such that:
   (a) $\vec{\mathbf{u}} + \vec{\mathbf{v}}$ is another vector, i.e. another element in $\mathcal{S}$
   (b) There is a zero vector in $\mathcal{S}$, called $\vec{\mathbf{0}}$, such that for every vector, $\vec{\mathbf{u}} + \vec{\mathbf{0}} = \vec{\mathbf{0}} + \vec{\mathbf{u}} = \vec{\mathbf{u}}$
   (c) For every $\vec{\mathbf{u}}$ there is another vector (called $-\vec{\mathbf{u}}$) in $\mathcal{S}$ such that $\vec{\mathbf{u}} + (-\vec{\mathbf{u}}) = \vec{\mathbf{0}}$
   (d) For any three vectors, $\vec{\mathbf{u}} + (\vec{\mathbf{v}} + \vec{\mathbf{w}}) = (\vec{\mathbf{u}} + \vec{\mathbf{v}}) + \vec{\mathbf{w}}$
   (e) $\vec{\mathbf{u}} + \vec{\mathbf{v}} = \vec{\mathbf{v}} + \vec{\mathbf{u}}$

A set having an operation satisfying the first four of these properties is called a *group*. When the operation, in our case addition, is commutative it is called an *Abelian group*.

A scalar multiple of a vector is multiplication by a real (or complex) number. Scalar multiplication gives us another vector in the space, so for all real numbers α and any vector $\vec{\mathbf{u}}$ in $\mathcal{S}$,

$$\vec{\mathbf{w}} = \alpha\vec{\mathbf{u}}$$

is also in $\mathcal{S}$.

Addition and scalar multiplication satisfy distributive laws,

$$\begin{aligned} \alpha\left(\vec{\mathbf{u}} + \vec{\mathbf{v}}\right) &= \alpha\vec{\mathbf{u}} + \alpha\vec{\mathbf{v}} \\ (\alpha+\beta)\,\vec{\mathbf{u}} &= \alpha\vec{\mathbf{u}} + \beta\vec{\mathbf{u}} \end{aligned}$$

These properties mean that for any collection of vectors, we may take arbitrary linear combinations, and the result is always another vector:

$$\vec{\mathbf{w}} = \alpha_1\vec{\mathbf{u}}_1 + \alpha_2\vec{\mathbf{u}}_2 + \ldots + \alpha_n\vec{\mathbf{u}}_n$$

The terms in this sum may be written in any order and added with any groupings.

1.2 Linear independence

A collection of vectors $\{\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2, \ldots, \vec{\mathbf{u}}_n\}$ is *linearly dependent* if there exist nonzero constants such that

$$\alpha_1 \vec{\mathbf{u}}_1 + \alpha_2 \vec{\mathbf{u}}_2 + \ldots + \alpha_n \vec{\mathbf{u}}_n = 0$$

When such a nontrivial sum of vectors vanishes, at least one of the αs must be nonzero, and we can divide through to solve for one of the vectors in terms of the others. Thus, if α_1 is one of the nonvanishing scalars, we may write

$$\vec{\mathbf{u}}_1 = -\frac{1}{\alpha_2} \left(\alpha_2 \vec{\mathbf{u}}_2 + \ldots + \alpha_n \vec{\mathbf{u}}_n\right)$$

This means that many vectors may be written in terms of a smaller set, called a *basis*.

A collection of vectors is *independent* if it is not dependent. This means that for independent vectors, the only solution above is for all of the $\alpha_1, \alpha_2, \ldots$ to vanish. A *basis* is any maximal independent set of vectors. For example, in 3 dimensions we know that we may write any vector as a linear combination of any three independent vectors. Let $\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{w}}$ be any 3 independent vectors in 3-space (so that the only way to get $\alpha\vec{\mathbf{u}} + \beta\vec{\mathbf{v}} + \gamma\vec{\mathbf{w}} = 0$ is to have $\alpha = \beta = \gamma = 0$). Then if $\vec{\mathbf{x}}$ is any other 3-vector the set

$$\{\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{w}}, \vec{\mathbf{x}}\}$$

is dependent. There exist nonzero numbers $\alpha, \beta, \gamma, \mu$ such that

$$\alpha\vec{\mathbf{u}} + \beta\vec{\mathbf{v}} + \gamma\vec{\mathbf{w}} + \mu\vec{\mathbf{x}} = 0$$

and necessarily μ is nonzero since the first three vectors are independent. Therefore,

$$\vec{\mathbf{x}} = -\frac{1}{\mu} \left(\alpha\vec{\mathbf{u}} + \beta\vec{\mathbf{v}} + \gamma\vec{\mathbf{w}}\right)$$

proving that any other vector in the space may be written as a linear combination of $(\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{w}})$. A basis gives a very concise way of describing any vector–all we need are the three real numbers, $\left(-\frac{\alpha}{\mu}, -\frac{\beta}{\mu}, -\frac{\gamma}{\mu}\right)$. These are the components of $\vec{\mathbf{x}}$ in the $(\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{w}})$ basis.

1.3 Norm and inner product

Now we come to the important ideas of magnitude and direction. These may be summarized for 3-vectors by the dot product. Essentially, the dot product is a way (one of many!) of producing a real number from a pair of vectors. Such a mapping on a vector space is called an inner product.

To find the squared length of any 3-vector–its *norm*–we have

$$\|\vec{\mathbf{x}}\|^2 := \vec{\mathbf{x}} \cdot \vec{\mathbf{x}}$$

When we have such a mapping that *linearly* assigns a positive real number to each vector, it is called a norm. Linearity means that

$$\begin{aligned} \|\alpha\vec{\mathbf{x}}\|^2 : &= (\alpha\vec{\mathbf{x}}) \cdot (\alpha\vec{\mathbf{x}}) \\ &= \alpha^2 (\vec{\mathbf{x}} \cdot \vec{\mathbf{x}}) \end{aligned}$$

and

$$\|\vec{\mathbf{x}} + \vec{\mathbf{y}}\|^2 : = (\vec{\mathbf{x}} + \vec{\mathbf{y}}) \cdot (\vec{\mathbf{x}} + \vec{\mathbf{y}})$$

The distributive property and symmetry of the dot product allow us to write this as

$$\begin{aligned} \|\vec{\mathbf{x}} + \vec{\mathbf{y}}\|^2 &= \vec{\mathbf{x}} \cdot \vec{\mathbf{x}} + \vec{\mathbf{y}} \cdot \vec{\mathbf{x}} + \vec{\mathbf{x}} \cdot \vec{\mathbf{y}} + \vec{\mathbf{y}} \cdot \vec{\mathbf{y}} \\ &= \|\vec{\mathbf{x}}\|^2 + 2\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} + \|\vec{\mathbf{y}}\|^2 \end{aligned}$$

and therefore

$$\begin{aligned} \|\vec{\mathbf{x}} + \vec{\mathbf{y}}\|^2 &= \vec{\mathbf{x}} \cdot \vec{\mathbf{x}} + \vec{\mathbf{y}} \cdot \vec{\mathbf{x}} + \vec{\mathbf{x}} \cdot \vec{\mathbf{y}} + \vec{\mathbf{y}} \cdot \vec{\mathbf{y}} \\ \vec{\mathbf{x}} \cdot \vec{\mathbf{y}} &= \frac{1}{2}\left(\|\vec{\mathbf{x}} + \vec{\mathbf{y}}\|^2 - \|\vec{\mathbf{x}}\|^2 - \|\vec{\mathbf{y}}\|^2\right) \end{aligned}$$

Conversely, if we were to take this as a definition of $\vec{\mathbf{x}} \cdot \vec{\mathbf{y}}$, it is automatically symmetric, $\vec{\mathbf{x}} \cdot \vec{\mathbf{y}} = \vec{\mathbf{y}} \cdot \vec{\mathbf{x}}$ and is defined in terms of vector length. This allows us to define an *inner product* between pairs of vectors once we have the squared *norm* for individual vectors.

A word of caution: there are other sorts of norms and vector spaces that do not automatically give an inner product.

1.4 An orthonormal basis

Using the $(\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{w}})$ basis for 3-space, we may write any vector as a linear combination,

$$\vec{\mathbf{x}} = x_u \vec{\mathbf{u}} + x_v \vec{\mathbf{v}} + x_w \vec{\mathbf{w}}$$

where (x_u, x_v, x_w) are the components of $\vec{\mathbf{x}}$. Then the squared norm of $\vec{\mathbf{x}}$ is

$$\begin{aligned} \|\vec{\mathbf{x}}\|^2 &= (x_u \vec{\mathbf{u}} + x_v \vec{\mathbf{v}} + x_w \vec{\mathbf{w}}) \cdot (x_u \vec{\mathbf{u}} + x_v \vec{\mathbf{v}} + x_w \vec{\mathbf{w}}) \\ &= x_u x_u \vec{\mathbf{u}} \cdot \vec{\mathbf{u}} + x_v x_u \vec{\mathbf{v}} \cdot \vec{\mathbf{u}} + x_w x_u \vec{\mathbf{w}} \cdot \vec{\mathbf{u}} + x_u x_v \vec{\mathbf{u}} \cdot \vec{\mathbf{v}} + x_v x_v \vec{\mathbf{v}} \cdot \vec{\mathbf{v}} \\ &\quad + x_w x_v \vec{\mathbf{w}} \cdot \vec{\mathbf{v}} + x_u x_w \vec{\mathbf{u}} \cdot \vec{\mathbf{w}} + x_v x_w \vec{\mathbf{v}} \cdot \vec{\mathbf{w}} + x_w x_w \vec{\mathbf{w}} \cdot \vec{\mathbf{w}} \end{aligned}$$

Even if we collect terms,

$$\begin{aligned} &= (x_u \vec{\mathbf{u}} + x_v \vec{\mathbf{v}} + x_w \vec{\mathbf{w}}) \cdot (x_u \vec{\mathbf{u}} + x_v \vec{\mathbf{v}} + x_w \vec{\mathbf{w}}) \\ \|\vec{\mathbf{x}}\|^2 &= x_u^2 \vec{\mathbf{u}} \cdot \vec{\mathbf{u}} + 2 x_u x_v \vec{\mathbf{u}} \cdot \vec{\mathbf{v}} + x_v^2 \vec{\mathbf{v}} \cdot \vec{\mathbf{v}} \\ &\quad + 2 x_u x_w \vec{\mathbf{u}} \cdot \vec{\mathbf{w}} + 2 x_v x_w \vec{\mathbf{v}} \cdot \vec{\mathbf{w}} + x_w^2 \vec{\mathbf{w}} \cdot \vec{\mathbf{w}} \end{aligned}$$

this is unnecessarily complicated. We may simplify it to the Pythagorean theorem if we require the basis vectors to be perpendicular or *orthogonal*,

$$\begin{aligned} \vec{\mathbf{u}} \cdot \vec{\mathbf{v}} &= 0 \\ \vec{\mathbf{u}} \cdot \vec{\mathbf{w}} &= 0 \\ \vec{\mathbf{w}} \cdot \vec{\mathbf{v}} &= 0 \end{aligned}$$

and *normalized* to unit length,

$$\begin{aligned} \vec{\mathbf{u}} \cdot \vec{\mathbf{u}} &= 1 \\ \vec{\mathbf{v}} \cdot \vec{\mathbf{v}} &= 1 \\ \vec{\mathbf{w}} \cdot \vec{\mathbf{w}} &= 1 \end{aligned}$$

There is a procedure for finding such a basis. The Gram-Schmidt procedure gives both properties at once, giving us an *orthonormal* basis. We start with any basis and choose one of the vectors, say $\vec{\mathbf{u}}$. If we divide $\vec{\mathbf{u}}$ by its norm it becomes a unit vector, denoted with a carat (^) instead of an arrow:

$$\hat{\mathbf{u}} = \frac{\vec{\mathbf{u}}}{\sqrt{\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}}}$$

Then

$$\begin{aligned} \hat{\mathbf{u}} \cdot \hat{\mathbf{u}} &= \frac{\vec{\mathbf{u}}}{\sqrt{\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}}} \cdot \frac{\vec{\mathbf{u}}}{\sqrt{\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}}} \\ &= \frac{\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}}{\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}} \\ &= 1 \end{aligned}$$

The vector $\hat{\mathbf{u}}$ is called a unit vector. It is the first vector in our orthonormal set. Now choose another vector from the basis, say $\vec{\mathbf{v}}$. We may always divide $\vec{\mathbf{v}}$ into parts parallel and perpendicular to $\hat{\mathbf{u}}$. The component of $\vec{\mathbf{v}}$ along $\hat{\mathbf{u}}$ is just

$$v_u = \hat{\mathbf{u}} \cdot \vec{\mathbf{v}}$$

so the part of $\vec{\mathbf{v}}$ parallel to $\vec{\mathbf{v}}$ is

$$\vec{\mathbf{v}}_\parallel = (\hat{\mathbf{u}} \cdot \vec{\mathbf{v}})\,\hat{\mathbf{u}} = v_u \hat{\mathbf{u}}$$

The rest of $\vec{\mathbf{v}}$ is given by the difference

$$\begin{aligned} \vec{\mathbf{v}}_\perp &= \vec{\mathbf{v}} - \vec{\mathbf{v}}_\parallel \\ &= \vec{\mathbf{v}} - (\hat{\mathbf{u}} \cdot \vec{\mathbf{v}})\,\hat{\mathbf{u}} \end{aligned}$$

and this gives the direction of our second basis vector. We only need to give it unit norm,

$$\hat{\mathbf{v}}_\perp = \frac{\vec{\mathbf{v}} - (\hat{\mathbf{u}} \cdot \vec{\mathbf{v}})\,\hat{\mathbf{u}}}{\|\vec{\mathbf{v}} - (\hat{\mathbf{u}} \cdot \vec{\mathbf{v}})\,\hat{\mathbf{u}}\|}$$

where the denominator is

$$\begin{aligned} \|\vec{\mathbf{v}} - (\hat{\mathbf{u}} \cdot \vec{\mathbf{v}})\,\hat{\mathbf{u}}\| &= \sqrt{(\vec{\mathbf{v}} - (\hat{\mathbf{u}} \cdot \vec{\mathbf{v}})\,\hat{\mathbf{u}}) \cdot (\vec{\mathbf{v}} - (\hat{\mathbf{u}} \cdot \vec{\mathbf{v}})\,\hat{\mathbf{u}})} \\ &= \sqrt{\vec{\mathbf{v}} \cdot \vec{\mathbf{v}} - 2\,(\hat{\mathbf{u}} \cdot \vec{\mathbf{v}})\,\hat{\mathbf{u}} \cdot \vec{\mathbf{v}} + (\hat{\mathbf{u}} \cdot \vec{\mathbf{v}})^2} \end{aligned}$$

Our basis now has two orthonormal vectors, $(\hat{\mathbf{u}}, \hat{\mathbf{v}}_\perp)$.

Exercise: Find a third unit vector orthogonal to both $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}_\perp$.

Given an orthonormal basis, we might as well align our coordinate axes with them and call them simpler names, such as $\left(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\right)$. Then every 3-vector may be written as

$$\vec{\mathbf{v}} = v_x \hat{\mathbf{i}} + v_y \hat{\mathbf{j}} + v_z \hat{\mathbf{k}}$$

and represented by the triple (v_x, v_y, v_z). The squared norm of $\vec{\mathbf{v}}$ is now given by the Pythagorean form,

$$\|\vec{\mathbf{v}}\|^2 = v_x^2 + v_y^2 + v_z^2$$

1.5 A note on notation

It is often useful to have a more general notation for an orthonormal basis. If we introduce a label, (i), that takes values $1, 2, 3$ and let

$$\begin{aligned} \hat{\mathbf{e}}_1 &= \hat{\mathbf{i}} \\ \hat{\mathbf{e}}_2 &= \hat{\mathbf{j}} \\ \hat{\mathbf{e}}_3 &= \hat{\mathbf{k}} \end{aligned}$$

then orthonormality may be expressed as

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$$

where δ_{ij} is the *Euclidean metric*, defined by

$$\begin{aligned} \delta_{11} &= \delta_{22} = \delta_{33} = 1 \\ \delta_{12} &= \delta_{23} = \delta_{31} = 0 \\ \delta_{21} &= \delta_{32} = \delta_{13} = 0 \end{aligned}$$

or more simply as the matrix

$$\delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Notice that $\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$ is actually nine different equations in which each of the labels i, j may take any of the values $1, 2, 3$.

Now, if we write the components of any vector $\vec{\mathbf{v}}$ as $\left(v^1, v^2, v^3\right)$ then expanding in the basis is a simple sum,

$$\vec{\mathbf{v}} = \sum_{i=1}^{3} v^i \hat{\mathbf{e}}_i$$

This notation allows easy generalization to vector spaces of any dimension. To write an n-dimensional vector, we just change the upper limit on the sum,

$$\vec{\mathbf{v}} = \sum_{i=1}^{n} v^i \hat{\mathbf{e}}_i$$

The orthonormality of the basis may still be written as $\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$ but this now expresses n^2 different equations as i and j each take values from 1 to n.

Finally, notice that we may use this label notation to compute the inner product. Using the distributive laws,

$$\begin{aligned} \vec{\mathbf{u}} \cdot \vec{\mathbf{v}} &= \left(\sum_{i=1}^{n} u^i \hat{\mathbf{e}}_i\right) \cdot \left(\sum_{j=1}^{n} v^j \hat{\mathbf{e}}_j\right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} u^i \hat{\mathbf{e}}_i \cdot v^j \hat{\mathbf{e}}_j \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} u^i v^j \left(\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j\right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} u^i v^j \delta_{ij} \end{aligned}$$

Now, since δ_{ij} vanishes unless $i = j$, this becomes

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} = u^1 v^1 + u^2 v^2 + \ldots + + u^n v^n$$

2 Function spaces

2.1 Taylor series

We know that many functions may be written in a Taylor series, including solutions to the wave equation. For a traveling wave,

$$q\left(x, t\right) = f\left(x - ct\right)$$

for example, we may expand $f\left(u\right)$ as about $u = 0$ as

$$f\left(u\right) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}\left(0\right) u^n$$

where $f^{(n)}(x) = \frac{d^n f}{dx^n}$. Now, with $u = x - ct$ we have

$$q(x,t) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0)(x-ct)^n$$

that is, the displacement q is written as a power series.

2.1.1 Power series as vectors

If we have two power series,

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0)\, x^n \\ g(x) &= \sum_{n=0}^{\infty} \frac{1}{n!} g^{(n)}(0)\, x^n \end{aligned}$$

we may take a linear combination of them to get a third,

$$\begin{aligned} h(x) &= \alpha f(x) + \beta g(x) \\ &= \alpha \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0)\, x^n + \beta \sum_{n=0}^{\infty} \frac{1}{n!} g^{(n)}(0)\, x^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\alpha f^{(n)}(0) + \beta g^{(n)}(0)\right) x^n \end{aligned}$$

If we think of $f(x)$ and $g(x)$ as countably infinite vectors,

$$\begin{aligned} f &\iff \left(f^{(0)}(0), f^{(1)}(0), f^{(2)}(0), \ldots\right) \\ g &\iff \left(g^{(0)}(0), g^{(1)}(0), g^{(2)}(0), \ldots\right) \end{aligned}$$

then this is exactly the way the components add, so that the n^{th} component of $h(x) = \alpha f(x) + \beta g(x)$ becomes

$$h^{(n)}(0) = \alpha f^{(n)}(0) + \beta g^{(n)}(0)$$

In fact, most of the properties we usually expect of vectors apply to functions. We can even define a squared norm such as

$$\|f\|^2 = \int_0^L |f(x)|^2\, dx$$

This choice is called the L^2 norm, and we will see that it generalizes the Pythagorean theorem for functions defined on the interval $[0, L]$.

If $f(x)$ and $g(x)$ are complex valued functions (for example, e^{ix}) then the norm is taken with the complex conjugate,

$$\langle f, g\rangle := \int_0^1 f^*(x)\, g(x)\, dx$$

This norm is used in quantum mechanics to find the probability of finding a particle in a volume V:

$$\|\psi\|^2 = \int_V \psi^*(\mathbf{x})\, \psi(\mathbf{x})\, d^3x$$

There are alternative norms. For polynomials on the double unit interval, it is more convenient to normalize by setting

$$f(1) = 1$$

We will use this norm below.

2.1.2 Inner products

We will want to know when two of these infinite dimensional vectors are orthogonal, but for orthogonality we need to know what we mean by the "dot product" of two functions.

Generalizing the result we found for 3-vectors,

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} = \frac{1}{2}\left(\|\vec{\mathbf{u}} + \vec{\mathbf{v}}\|^2 - \|\vec{\mathbf{u}}\|^2 - \|\vec{\mathbf{v}}\|^2\right)$$

we can define an inner product for a function space from its norm. The L^2 norm allows us to do this. Consider the inner product of $h(x) = f(x) + g(x)$ in the L^2 norm:

$$\begin{aligned}
\|f+g\|^2 &= \int_0^1 (f(x) + g(x))^2 \, dx \\
&= \int_0^1 f(x)^2 + g(x)^2 + 2f(x)g(x) \, dx \\
&= \|f\|^2 + \|g\|^2 + 2\int_0^1 f(x)g(x) \, dx \\
\int_0^1 f(x)g(x) \, dx &= \frac{1}{2}\left(\|f+g\|^2 - \|f\|^2 - \|g\|^2\right)
\end{aligned}$$

We define this to be our inner product:

$$\langle f, g \rangle := \int_0^1 f(x)g(x) \, dx$$

This method does not work for every norm. For example, if we had defined the L^4 norm

$$\|f\|^4 = \int_0^L |f(x)|^4 \, dx$$

then

$$\begin{aligned}
\|f+g\|^4 &= \int_0^1 (f(x) + g(x))^4 \, dx \\
&= \int_0^1 \left(f^4 + 4f^3g + 6f^2g^2 + 4fg^3 + g^4\right) dx \\
&= \|f\|^4 + \|g\|^4 + \int_0^1 \left(4f^3g + 6f^2g^2 + 4fg^3\right) dx
\end{aligned}$$

so although we could define

$$
\begin{aligned}
\langle f, g\rangle &= \|f+g\|^4 - \|f\|^4 - \|g\|^4 \\
&= \int_0^1 \left(4f^3 g + 6f^2 g^2 + 4fg^3\right) dx
\end{aligned}
$$

this expression has little to recommend it. With more general definitions of the norm, it becomes impossible to discern a meaningful inner product.

2.1.3 An orthonormal basis for a function space

For the vector space of power series, the basis $1, x, x^2, \ldots$ is not orthogonal. For example, on the interval $[-1,1]$,,

$$\int_{-1}^{1} \left(x^2\right) \cdot \left(x^4\right) dx = \frac{1}{7}\, x^7\Big|_{-1}^{1} = \frac{2}{7}$$

so that x^2 and x^4 are not orthogonal. Therefore, let us apply Gram-Schmidt orthogonalization to find an orthogonal basis. For normalization we will choose to set $f(1) = 1$.

To find such a basis for a power series, it serves future purposes if we consider functions on the unit interval, $[-1,1]$. Then we begin with the simplest polynomial, a constant

$$P_0(x) = 1$$

and this already satisfies $P_0(1) = 1$. This is what we mean by a unit vector on this function space.

Next consider a linear polynomial,

$$P_1(x) = ax + b$$

We first make it orthogonal to φ_0,

$$
\begin{aligned}
0 &= \langle P_0, P_1\rangle \\
&= \int_{-1}^{1} 1 \cdot (ax+b)\, dx \\
&= \left(\frac{1}{2}ax^2 + bx\right)\Big|_{-1}^{1} \\
&= \left(\frac{1}{2}a + b\right) - \left(\frac{1}{2}a - b\right) \\
&= 2b
\end{aligned}
$$

Evaluating at $x = 1$, we require

$$1 = P_1(1) = a$$

so we need $a = 1$ to give the final form

$$P_1(x) = x$$

For P_2, set $P_2(x) = ax^2 + bx + c$, and require

$$
\begin{aligned}
\langle P_0, P_2\rangle &= 0 \\
\langle P_1, P_2\rangle &= 0 \\
\|P_2\|^2 &= 1
\end{aligned}
$$

For the first

$$\begin{aligned}
0 &= \langle P_0, P_2 \rangle \\
&= \int_{-1}^{1} 1 \cdot \left(ax^2 + bx + c\right) dx \\
&= \left(\frac{1}{3}ax^3 + \frac{1}{2}bx^2 + cx\right)\Big|_{-1}^{1} \\
&= \frac{1}{3}a + \frac{1}{2}b + c - \left(-\frac{1}{3}a + \frac{1}{2}b - c\right) \\
&= \frac{2}{3}a + 2c
\end{aligned}$$

so that

$$c = -\frac{1}{3}a$$

Next,

$$\begin{aligned}
0 &= \langle P_1, P_2 \rangle \\
&= \int_{-1}^{1} (2x-1) \cdot \left(ax^2 + bx + c\right) dx \\
&= \int_{-1}^{1} \left(2ax^3 + 2bx^2 + 2cx - ax^2 - bx - c\right) dx \\
&= \int_{-1}^{1} \left(2ax^3 + (2b-a)\,x^2 + (2c-b)\,x - c\right) dx \\
&= \frac{1}{2}a + \frac{2}{3}b - \frac{1}{3}a - \frac{1}{2}b - \left(\frac{1}{4}2a - \frac{1}{3}(2b-a) + \frac{1}{2}(2c-b) + c\right) \\
&= \frac{1}{2}a + \frac{2}{3}b - \frac{1}{3}a - \frac{1}{2}b - \frac{1}{2}a + \frac{1}{3}(2b-a) - \frac{1}{2}(2c-b) - c \\
&= -\frac{1}{3}a - \frac{1}{3}a + \frac{2}{3}b - \frac{1}{2}b + \frac{2}{3}b + \frac{1}{2}b - 2c \\
&= -\frac{2}{3}a + \frac{4}{3}b - 2c \\
&= -\frac{2}{3}a + \frac{4}{3}b + \frac{2}{3}a \\
b &= 0
\end{aligned}$$

Now we have

$$\begin{aligned}
P_2(x) &= a\left(x^2 - \frac{1}{3}\right) \\
P_2(1) &= a\left(\frac{2}{3}\right)
\end{aligned}$$

so we choose $a = \frac{3}{2}$. Thus

$$P_2(x) = \frac{1}{2}\left(3x^2 - 1\right)$$

The polynomials generated in this way are called the *Legendre polynomials*. They form an orthonormal basis for power series on the interval $[-1,1]$. (We will discover them again when we study the Laplace equation in spherical coordinates–the angle θ varies from 0 to π, so that $\cos\theta$ varies from -1 to 1 and we can use them so express any function of θ.)

Exercise: Find $P_3(x)$

Because these polynomials are orthogonal, inner products only depend on the their norms. There is a general formula for the L^2 norm of P_n. For the three we have,

$$\begin{aligned}
\langle P_0, P_0\rangle &= \int_{-1}^{1} 1\cdot 1dx \\
&= 2 \\
\langle P_1, P_1\rangle &= \int_{-1}^{1} x\cdot xdx \\
&= \frac{2}{3} \\
\langle P_2, P_2\rangle &= \int_{-1}^{1} \frac{1}{2}\left(3x^2-1\right)\cdot\frac{1}{2}\left(3x^2-1\right)dx \\
&= \frac{1}{4}\int_{-1}^{1}\left(9x^4-6x^2+1\right)dx \\
&= \frac{1}{4}\left(\frac{9}{5}x^5-\frac{6}{3}x^3+x\right)\Big|_{-1}^{1} \\
&= \frac{2}{4}\left(\frac{9}{5}-\frac{6}{3}+1\right) \\
&= \frac{2}{5}
\end{aligned}$$

In general it can be shown that the L^2 norm of $P_n(x)$ is

$$\langle P_n, P_n\rangle = \frac{2}{2n+1}$$

2.1.4 Expanding functions in Legendre polynomials

Because of the orthonormality, it often turns out to be convenient to write a power series in terms of Legendre polynomials instead of simply powers of x,

$$f(x) = \sum_{n=0}^{\infty} a_n\varphi_n(x)$$

As an example, consider the polynomial

$$f(x) = a + bx + cx^2$$

We may rewrite this by first replacing x^2 with a multiple of $P_2(x)$,

$$\begin{aligned}
P_2(x) &= \frac{1}{2}\left(3x^2-1\right) \\
2P_2(x) &= 3x^2-1 \\
x^2 &= \frac{1}{3}\left(2P_2(x)+1\right)
\end{aligned}$$

Using this

$$\begin{aligned} f(x) &= a + bx + \frac{c}{3}\left(2P_2(x) + 1\right) \\ &= \left(a + \frac{c}{3}\right) + bx + \frac{2c}{3}P_2(x) \end{aligned}$$

Now replace x by $P_1(x)$ and write the constant as a multiple of $P_0(x) = 1$,

$$f(x) = \left(a + \frac{c}{3}\right)P_0(x) + bP_1(x) + \frac{2c}{3}P_2(x)$$

Computations become much easier in this form. For instance, if we want to compute the L^2 norm of f,

$$\|f(x)\|^2 = \int_{-1}^{1}\left(\left(a + \frac{c}{3}\right)P_0(x) + bP_1(x) + \frac{2c}{3}P_2(x)\right)\left(\left(a + \frac{c}{3}\right)P_0(x) + bP_1(x) + \frac{2c}{3}P_2(x)\right)$$

This would be complicated as a power series, giving us nine integrals to perform, but orthogonality eliminates all but three integrals, and we already know their values,

$$\begin{aligned} \|f(x)\|^2 &= \int_{-1}^{1}\left(\left(a + \frac{c}{3}\right)P_0(x) + bP_1(x) + \frac{2c}{3}P_2(x)\right)\left(\left(a + \frac{c}{3}\right)P_0(x) + bP_1(x) + \frac{2c}{3}P_2(x)\right) \\ &= \left(a + \frac{c}{3}\right)^2\langle P_0(x), P_0(x)\rangle + b^2\langle P_1(x), P_1(x)\rangle + \frac{4c}{9}\langle P_2(x), P_2(x)\rangle \\ &= \left(a + \frac{c}{3}\right)^2 2 + b^2\frac{2}{3} + \frac{4c}{9}\frac{2}{5} \\ &= 2a^2 + \frac{4}{3}ac + \frac{2}{9}c^2 + \frac{2}{3}b^2 + \frac{8}{45}c \end{aligned}$$

More generally, suppose we have two functions expanded in Legendre polynomials,

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} a_n P_n(x) \\ g(x) &= \sum_{m=0}^{\infty} b_m P_m(x) \end{aligned}$$

Then their L^2 inner product is

$$\begin{aligned} \langle f(x), g(x)\rangle &= \int_{-1}^{1} f(x)\,g(x)\,dx \\ &= \sum_{n=0}^{\infty} a_n \sum_{m=0}^{\infty} b_m \int_{-1}^{1} P_n(x)\,P_m(x)\,dx \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} a_n b_m \left(\frac{2}{2n+1}\delta_{nm}\right) \end{aligned}$$

and since δ_{mn} is zero except when $m = n$ we can do one of the sums, just changing m to n,

$$\begin{aligned} \langle f(x), g(x)\rangle &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} a_n b_m \frac{2}{2n+1}\delta_{nm} \\ &= \sum_{n=0}^{\infty} \frac{2a_n b_n}{2n+1} \end{aligned}$$

We've just performed an infinite number of integrals!