Examples Using Orthogonal Vectors

Simple Example Say you need to solve the equations

$$\begin{aligned} x_1 + x_2 + \quad x_3 + x_4 &= \quad y_1 \\ x_2 - x_2 - \quad x_3 + x_4 &= \quad y_2 \\ -x_1 + x_2 - \quad x_3 + x_4 &= \quad y_3 \\ -x_1 - x_2 + \quad x_3 + x_4 &= \quad y_4 \end{aligned}$$

for $x_1,\ x_2,\ x_2,\ x_4$. Rewrite this as

$$x_1\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix} + x_2\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix} + x_3\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix} + x_4\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix},$$

that is,

$$x_1V_1 + x_2V_2 + x_3V_3 + x_4V_4 = Y,$$

where the V_j and Y are the obvious vectors. The key observation is that these vectors V_j are orthogonal and have length $\|V_j\| = 2$. It is now simple to solve the equations. Taking the inner product of both sides with V_1 we get

$$x_1\langle V_1, V_1\rangle + x_2\langle V_2, V_1\rangle + x_3\langle V_3, V_1\rangle + x_4\langle V_4, V_1\rangle = \langle Y, V_1\rangle,$$

that is,

$$x_1\|V_1\|^2 + 0 + 0 + 0 = \langle Y, V_1\rangle, \qquad \text{so} \qquad x_1 = \tfrac{1}{4}\langle Y, V_1\rangle.$$

By the same procedure,

$$x_j = \tfrac{1}{4}\langle Y, V_j\rangle, \qquad j = 1, 2, 3, 4.$$

Not hard work at all.

While it may seem exotic (and lucky) that the vectors V_j were orthogonal, it turns out that this arises naturally – and frequently – in very important applications. For instance when Fourier series arise and in the analysis of large data sets..

Orthogonal Projection

Let V be an inner product space (that is, a linear space with an inner product) and let $\vec{v}_1, \vec{v}_2, \ldots,$ $\vec{v}_k$ be non-zero orthogonal vectors and let $\mathcal{S} \subset V$ be the subspace spanned by these $\vec{v}_j$'s. Given a vector $\vec{x} \in V$, we want to write

$$\vec{x} = \vec{v} + \vec{w}, \tag{1}$$

where $\vec{v} \in \mathcal{S}$ and $\vec{w} \perp \mathcal{S}$. We then call $\vec{v}$ the *orthogonal projection of $\vec{x}$ into $\mathcal{S}$* and often write $\vec{v} = P_{\mathcal{S}}\vec{x}$.

Because we know the $\vec{v}_j$ are an orthogonal basis for $\mathcal{S}$, then any vector $\vec{v} \in \mathcal{S}$ can be written as

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots a_k\vec{v}_k$$

so we can write $\vec{x}$ as

$$\vec{x} = (a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k) + \vec{w}, \tag{2}$$

where $\vec{w}$ is orthogonal to $\mathcal{S}$. This decomposes $\vec{x}$ as the sum of two orthogonal vectors, $\vec{v}$ in $\mathcal{S}$ and one, $\vec{w}$ orthogonal to $\mathcal{S}$. We often introduce the linear map $P_{\mathcal{S}}$ of orthogonal projection into $\mathcal{S}$

$$P_{\mathcal{S}}\vec{x} := \vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k.$$

If we write $\mathcal{S}^\perp$ for the orthogonal complement of $\mathcal{S}$, then $\vec{w} = P_{\mathcal{S}^\perp}\vec{x}$, so

$$\vec{x} = \vec{v} + \vec{w} = P_{\mathcal{S}}\vec{x} + P_{\mathcal{S}^\perp}\vec{x} = (a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k) + \vec{w}.$$

The problem is to find the coefficients a_j and the vector $\vec{w}$. Easy!

Taking the inner product of both sides of equation (2) with $\vec{v}_1$ we find hat $\langle\vec{x},\, \vec{v}_1\rangle = a_1\langle\vec{v}_1,\, \vec{v}_1\rangle$ and similarly for the other a_j's. Thus

$$a_j = \frac{\langle\vec{x},\, \vec{v}_j\rangle}{\|\vec{v}_j\|^2}, \tag{3}$$

so we now know the coefficients a_j in equation (2). We can now solve equation (2) for $\vec{w}$ and find

$$\vec{w} = \vec{x} - (a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k),$$

Since the $\vec{v}_j$'s and $\vec{w}$ are orthogonal, the Pythagorean theorem applied to (2) tells us that

$$\|\vec{x}\|^2 = |a_1|^2\|\vec{v}_1\|^2 + \cdots + |a_k|^2\|\vec{v}_k\|^2 + \|\vec{w}\|^2. \tag{4}$$

In particular,

$$\|\vec{w}\|^2 = \|\vec{x}\|^2 - \|P_S\vec{x}\|^2 = \|\vec{x}\|^2 - \left(|a_1|^2\|\vec{v}_1\|^2 + \cdots + |a_k|^2\|\vec{v}_k\|^2\right) \tag{5}$$

gives the *square of the distance from $\vec{x}$ to the subspace $\mathcal{S}$.*

Remark: There are two slightly different approaches to finding the distance from a point $\vec{x}$ to a subspace $\mathcal{S}$. In both approaches we end up computing

$$\text{Distance } = \|P_{\mathcal{S}^\perp}\vec{x}\|$$

METHOD 1 Find the orthogonal projection $\vec{v} = P_{\mathcal{S}}\vec{x}$. Then, as we found above, the orthogonal projection into $S^\perp$ is $\vec{w} = P_{\mathcal{S}^\perp}\vec{x} = \vec{x} - P_{\mathcal{S}}\vec{x}$.

METHOD 2 Directly compute the orthogonal projection into $\mathcal{S}^\perp$. For this approach, the first step is usually to find an orthogonal basis for $\mathcal{S}$ and then extend this as an orthogonal basis to the $S^\perp$. This usually involves far more computations – but there is one frequently occurring situation where it is very easy: when the dimension of $\mathcal{S}^\perp$ is one.

Here is an Example. Let $\mathcal{S}$ be the plane in $\mathbb{R}^3$ where $ax_1 + bx_2 + cx_3 = 0$. If we let $\vec{N} = (a,\, b,\, c)$, then the equation for the plane is simply $\langle\vec{x},\, \vec{N}\rangle = 0$. Thus $\vec{N}$ is an orthogonal basis for $\mathcal{S}^\perp$ – and one never need to even find an orthogonal basis for $\mathcal{S}$ itself. The orthogonal projection of $\vec{x}$ into $\mathcal{S}^\perp$ is then simply

$$\vec{w} = \frac{\langle\vec{x},\, \vec{N}\rangle}{\|\vec{N}\|^2}\vec{N},$$

so the length of this vector $\vec{w}$, $\frac{|\langle\vec{x},\, \vec{N}\rangle|}{\|\vec{N}\|}$, gives the distance from $\vec{x}$ to $\mathcal{S}$.

Example In $\mathbb{R}^4$, let the subspace $\mathcal{S}$ be the span of the vectors $\vec{v}_1 := (1, 1, -1, -1)$ and $\vec{v}_2 := (1, 1, 1, 1)$.

a) Find the orthogonal projection of $\vec{x} := (1, 2, 3, 4)$ into $\mathcal{S}$.

b) Find the distance from $\vec{x}$ to the plane $\mathcal{S}$.

Solution: (a) Note that the vectors $\vec{v}_1$ and $\vec{v}_2$ are an *orthogonal* basis for $\mathcal{S}$. We want to write

$$\vec{x} = a_1\vec{v}_1 + a_2\vec{v}_2 + \vec{w}, \tag{6}$$

where $\vec{w} \perp \mathcal{S}$. Then the orthogonal projection of $\vec{x}$ into $\mathcal{S}$ will be

$$P_{\mathcal{S}}\vec{x} = a_1\vec{v}_1 + a_2\vec{v}_2,$$

By the general strategy use above, to find a_1 take the inner product of both sides of equation(6) with $\vec{v}_1$. Because $\vec{v}_1$ is orthogonal to both $\vec{v}_2$ and w, we obtain

$$\langle\vec{x},\, \vec{v}_1\rangle = a_1\langle\vec{v}_1,\, \vec{v}_1\rangle \qquad \text{so} \qquad a_1 = \frac{\langle\vec{x},\, \vec{v}_1\rangle}{\|\vec{v}_1^2\|} = \frac{-4}{4} = -1.$$

Similarly,

$$a_2 = \frac{\langle\vec{x},\, \vec{v}_2\rangle}{\|\vec{v}_2^2\|} = \frac{10}{4} = \frac{5}{2}.$$

Using these values in equation (6) we find the projection of $\vec{x}$ into $\mathcal{S}$ is

$$P_{\mathcal{S}}\vec{x} = -\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} + \frac{5}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 3 \\ 3 \\ 7 \\ 7 \end{pmatrix}$$

and the projection of $\vec{x}$ orthogonal to $\mathcal{S}$ is

$$\vec{w} = P_{\mathcal{S}^\perp}\vec{x} = \vec{x} - P_{\mathcal{S}}\vec{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 3 \\ 3 \\ 7 \\ 7 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

As a check, this $\vec{w}$ is clearly orthogonal to $\mathcal{S}$.

(b) Finally, using equation (5), the distance from the point $\vec{x}$ to this subspace $\mathcal{S}$ is $\|\vec{w}\| = 1$.

Exercises

1. Find the distance between the point $\vec{x} = (1, 2, -3, 0) \in \mathbb{R}^4$ and the subspace of points $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ that satisfy $x_1 - x_2 + x_3 + 2x_4 = 0$.

2. Find the distance between the hyperplane of points $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ that satisfy $x_1 - x_2 + x_3 + 2x_4 = 2$ and the origin.

3. In $\mathbb{R}^5$, let $\mathcal{S}$ be the subspace spanned by the vectors $\vec{v}_1 = (1,\, 1,\, -1,\, 0,\, -1)$ and $\vec{v}_2 = (1,\, 1,\, 1,\, 0,\, 1)$. Find the orthogonal projection of $\vec{x} = (1,\, 0,\, 0,\, 1,\, -1)$ into $\mathcal{S}$ and compute the distance from $\vec{x}$ to $\mathcal{S}$.

4. Find an orthogonal basis for the subspace of $\mathbb{R}^4$ spanned by $\vec{u}_1 = (1, 1, 0, 0)$ and $\vec{u}_2 = (0, 1, 1, 0)$

5. Find a vector in $\mathbb{R}^4$ that is orthogonal to the subspace spanned by $\vec{u}_1 = (1,1,0,0)$ and $\vec{u}_2 = (0,1,1,0)$.

6. Find an orthogonal basis for the subspace of $\mathbb{R}^4$ spanned by $\vec{u}_1 = (1,1,0,0)$, $\vec{u}_2 = (0,1,1,0)$, and $\vec{u}_3 = (0,0,1,1)$.

7. Find an orthonormal basis for the sub-apace of $\mathbb{R}^4$ determined by $x_1 - x_2 + x_3 - 2x_4 = 0$.

8. Find a vector that is orthogonal to the above subspace.

Example: Fourier Series

The essential point of this next example is that the formalism using the inner product that we have just developed in $\mathbb{R}^n$ is immediately applicable in a much more general setting – with wide and important applications. We use geometric intuition from $\mathbb{R}^n$ to guide us through related ideas in infinite dimensional function spaces.

Here our linear space is $L_2(-\pi,\, \pi)$ with a standard (real) inner product

$$\langle f,\, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\, dx$$

and are using the linear space

$$\mathcal{T}_N = \operatorname{span}\{1, \cos x, \cos 2x, \ldots, \cos Nx,\ \sin x, \ldots, \sin Nx\}.$$

An *orthonormal* basis is:

$$e_0 := \frac{1}{\sqrt{2\pi}}, \quad e_1 := \frac{\cos x}{\sqrt{\pi}}, \ldots,\ e_N := \frac{\cos Nx}{\sqrt{\pi}}, \quad \epsilon_1 := \frac{\sin x}{\sqrt{\pi}}, \ldots,\ \epsilon_N := \frac{\sin Nx}{\sqrt{\pi}}.$$

We want to find the projection of a given function $f(x)$ into $\mathcal{T}_N$, that is, write

$$f(x) = a_0 e_0 + (a_1 e_1 + \cdots + a_N e_N) + (b_1 \epsilon_1 + \cdots + b_N \epsilon_N) + h_N, \tag{7}$$

where the "error," h_N, is orthogonal to $\mathcal{T}_N$. This problem is *exactly* of the form of equation (2). Thus we can use all the results we obtained there.

First, we have a formula for the coefficients. This is a bit simpler here than the formula in equation (3) since $e_k(x)$ and $\epsilon_k(x)$ have $\|e_k\| = \|\epsilon_k\| = 1$.

$$a_0 = \langle f,\, e_0\rangle, \qquad \text{while} \qquad a_k = \langle f,\, e_k\rangle, \qquad b_k = \langle f,\, \epsilon_k\rangle, \qquad j = .2, 3, \ldots.$$

Using the explicit formulas for the e_k and ϵ_k we have

$$f(x) = a_0 \frac{1}{\sqrt{2\pi}} + \sum_{k=1}^{N} \left[a_k \frac{\cos kx}{\sqrt{\pi}}\, dx + b_k \frac{\sin kx}{\sqrt{\pi}} \right] + h_N(x), \tag{8}$$

where, as above, h_N is orthogonal to $\mathcal{T}_N$. Series of this form are called *Fourier Series*. They are a vital ingredient in today's world, including quantum mechanics, medical imaging and your cell phone.

For the coefficients we have

$$a_0 = \int_{-\pi}^{\pi} f(x)\frac{1}{\sqrt{2\pi}}\,dx, \qquad a_k = \int_{-\pi}^{\pi} f(x)\frac{\cos kx}{\sqrt{\pi}}\,dx, \qquad b_k = \int_{-\pi}^{\pi} f(x)\frac{\sin kx}{\sqrt{\pi}}\,dx. \tag{9}$$

These coefficients incorporate that $h_N(x)$ is orthogonal to $\mathcal{T}_N$. To summarize,

$$f(x) = P_{\mathcal{T}_N} f(x) + h_N(x) = a_0\frac{1}{\sqrt{2\pi}} + \sum_{k=1}^{N}\left[a_k\frac{\cos kx}{\sqrt{\pi}} + b_k\frac{\sin kx}{\sqrt{\pi}}\right] + h_N(x)$$

Of course, one hopes that $\lim_{N\to\infty} \|h_N\|_{L_2(-\pi.\pi)} = 0$. It is true for essentially all functions, certainly for all piecewise continuous functions f. The above series is called the *Fourier Series of* $f(x)$.

The Pythagorean formula (4) gives

$$\|f\|^2_{L_2(-\pi.\pi)} = |a_0|^2 + \sum_{k=1}^{N}\Big(|a_k|^2 + |b_k|^2\Big) + \|h_N\|^2_{L_2(-\pi.\pi)}. \tag{10}$$

Privately, I call equation (10) the "Pythagorean Theorem for Adults".

Explicit Example: Fourier Series of a Square Wave

Consider the function $f(x) = \begin{cases} -1 & \text{if } -\pi < x \le 0 \\ \;\;1 & \text{if } \;\; 0 < x \le \pi \end{cases}$

We use equation (9) to compute the Fourier coefficients a_k and b_k.
Since this $f(x)$ is an odd function, then $f(x)\cos kx$ is also an odd function so $a_k = 0,\ k = 0,\,1,\ldots$.
Similarly, using that $f(x)\sin kx$ is an even function, we have

$$b_k = \frac{1}{\sqrt{\pi}}\left[\int_{-\pi}^{0}(-1)\sin kx\,dx + \int_{0}^{\pi}(+1)\sin kx\,dx\right] = \frac{2}{\sqrt{\pi}}\int_0^{\pi}\sin kx\,dx.$$

But

$$\int_0^{\pi}\sin kx\,dx = \frac{-\cos k\pi + 1}{k} = \begin{cases} 0 & \text{if } k \text{ is even} \\ \frac{2}{k} & \text{if } k \text{ is odd}\end{cases}.$$

Therefore

$$b_k = \begin{cases} 0 & \text{if } k \text{ is even} \\ \frac{4}{k\sqrt{\pi}} & \text{if } k \text{ is odd}\end{cases}.$$

We now substitute this into equation (8) and write $N = 2n+1$ to obtain the following Fourier Series of a square wave:

$$f(x) = \frac{4}{\pi}\left[\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots + \frac{\sin(2n+1)x}{2n+1}\right] + h_{2n+1}(x).$$

Here is a graph showing how the terms in this series approximate a square wave:
`http://www.math.upenn.edu/~kazdan/312S14/notes/Fourier-SquareWave.gif`
[From Wolfram *MathWorld*]

Finally we record the Pythagorean formula (10). Since in our case $f(x)^2 = 1$, then $\int_{-\pi}^{\pi} f(x)^2\,dx = 2\pi$ and equation (10) give

$$2\pi = \frac{16}{\pi}\left[1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots + \frac{1}{(2n+1)^2}\right] + \|h_{2n+1}\|^2.$$

With some work one can show that $\lim_{n\to\infty}\|h_{2n+1}\| = 0$. This yields the surprising formula

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots \qquad (11)$$

Subtracting

$$\frac{1}{4}\left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots\right) = \frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{6^2} + \frac{1}{8^2} + \cdots$$

from

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots = \left[1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right] + \left[\frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{6^2} + \cdots\right]$$

and using equation (11), by a simple computation we obtain

$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots$$

It is amazing that identifies like these are rather immediate consequences of the Pythagorean Theorem. Not at all obvious.

[Last revised: February 27, 2014]