2
Gram-Schmidt Orthogonalization

One very useful characteristic of Hermitian operators is that the non-degenerate eigenfunctions (that is, eigenfunctions that don't share an eigenvalue with other eigenfunctions) of a Hermitian operator are guaranteed to be orthogonal to one another. And as mentioned in Section 2.3 of *A Student's Guide to the Schrödinger Equation*, even in the case of degenerate eigenfunctions (which share eigenvalues with other eigenfunctions and may not be orthogonal), those eigenfunctions can be used to construct a set of orthogonal eigenfunctions.

This document describes one approach for doing that construction: the Gram-Schmidt orthogonalization procedure. Using this procedure, you'll be able to construct a set of orthogonal vectors or functions from a set of non-orthogonal vectors or functions. After that description, you'll find a list of references with more details of the Gram-Schmidt procedure applied to both vectors and functions. The concept of Gram-Schmidt orthogonalization is straightforward, involving the projection of one of the vectors onto another (to find out how much of that vector "lies along" the other) and then subtracting that projection from the original vector, leaving on the part that doesn't "lie along" the original vector.

So in the case of two vectors such as $\vec{A}$ and $\vec{B}$, start by projecting vector $\vec{B}$ onto vector $\vec{A}$ (you could equally well project vector $\vec{A}$ onto vector $\vec{B}$) and then turn that (scalar) projection into a vector in the direction of vector $\vec{A}$. So now you've got a vector in the direction of $\vec{A}$ but with magnitude equal to the "amount" of vector $\vec{B}$ along the direction of vector $\vec{A}$. So if you then subtract that new vector from vector $\vec{B}$, you'll be left with the part of $\vec{B}$ that doesn't lie at all along the direction of $\vec{A}$. In other words, you'll have made a vector that's perpendicular to $\vec{A}$, which is exactly what you were trying to do. So vector $\vec{A}$ hasn't changed, and you've turned vector $\vec{B}$ into a vector that's perpendicular to $\vec{A}$.

Here's how the process looks in equations for two vectors $\vec{A}$ and $\vec{B}$:

$$\vec{A}_{new} = \vec{A}_{orig}$$

$$\vec{B}_{new} = \vec{B}_{orig} - \textit{ part of } \vec{B}_{orig} \textit{ lying along } \vec{A}_{new} \; = \; \vec{B}_{orig} - \frac{\vec{B}_{orig} \circ \vec{A}_{new}}{\vec{A}_{new} \circ \vec{A}_{new}} \vec{A}_{new}.$$

If you want to make your new set of vectors orthonormal (that is, not only orthogonal but also with unit magnitude), just divide each of them by its magnitude.

You may be thinking,"Ok, but what if I have a set of non-orthogonal three-dimensional vectors, or ten-dimensional abstract vectors, or continuous functions?" Happily, the Gram-Schmidt procedure works for those cases as well, since you can just keep using the same trick (projecting and subtracting off). For example, if you have three potentially non-orthogonal vectors $\vec{A}$, $\vec{B}$, and $\vec{C}$ in a three-dimensional vector space, use the procedure described above to find the part of $\vec{B}$ that's perpendicular to $\vec{A}$, and then find the part of $\vec{C}$ that's perpendicular to both $\vec{A}$ and $\vec{B}$ (by projecting $\vec{C}$ onto both $\vec{A}$ and $\vec{B}$ and then subtracting off those parts of $\vec{C}$).

Consider the following three vectors expressed in the three-dimensional Cartesian coordinate system:

$$\vec{A}_{orig} = \begin{pmatrix} A_x \\ A_y \\ A_y \end{pmatrix} \qquad \vec{B}_{orig} = \begin{pmatrix} B_x \\ B_y \\ B_y \end{pmatrix} \qquad \vec{C}_{orig} = \begin{pmatrix} C_x \\ C_y \\ C_z \end{pmatrix}.$$

Using the Gram-Schmidt procedure to find three orthogonal vectors ($\vec{A}_{new}$, $\vec{B}_{new}$, and $\vec{C}_{new}$) based on these three vectors looks like this:

$$\vec{A}_{new} = \vec{A}_{orig} = \begin{pmatrix} A_x \\ A_y \\ A_y \end{pmatrix}$$

$$\vec{B}_{new} = \vec{B}_{orig} - \textit{ part of } \vec{B}_{orig} \textit{ lying along } \vec{A}_{new} \; = \; \vec{B}_{orig} - \frac{\vec{B}_{orig} \circ \vec{A}_{new}}{\vec{A}_{new} \circ \vec{A}_{new}} \vec{A}_{new}$$

$$\begin{aligned} \vec{C}_{new} &= \vec{C}_{orig} - \textit{ part of } \vec{C}_{orig} \textit{ lying along } \vec{A}_{new} - \textit{ part of } \vec{C}_{orig} \textit{ lying along } \vec{B}_{new} \\ &= \vec{C}_{orig} - \frac{\vec{C}_{orig} \circ \vec{A}_{new}}{\vec{A}_{new} \circ \vec{A}_{new}} \vec{A}_{new} - \frac{\vec{C}_{orig} \circ \vec{B}_{new}}{\vec{B}_{new} \circ \vec{B}_{new}} \vec{B}_{new} \end{aligned}$$

in which the open circles represent the dot product.

To see this procedure in action, consider the following three-dimensional vectors expressed in the Cartesian coordinate system:

$$\vec{A} = 3\hat{\imath} \ - 2\hat{\jmath} + 4\hat{k} \qquad \vec{B} = \hat{\jmath} \ - 2\hat{k} \qquad \vec{C} = 2\hat{\imath} \ + 3\hat{\jmath} \ - \hat{k}.$$

Writing these vectors as column vectors makes them

$$\vec{A}_{orig} = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} \qquad \vec{B}_{orig} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} \qquad \vec{C}_{orig} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$$

and the Gram-Schmidt equations give

$$\vec{A}_{new} = \vec{A}_{orig} = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}$$

$$\vec{B}_{new} = \ \vec{B}_{orig} - \ \frac{\vec{B}_{orig} \ \circ \vec{A}_{new}}{\vec{A}_{new} \ \circ \vec{A}_{new}} \ \vec{A}_{new} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} - \frac{\begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} \circ \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}}{\begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} \circ \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} - \frac{\begin{pmatrix} 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}}{\begin{pmatrix} 3 & -2 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} - \frac{-10}{29} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} + \frac{1}{29} \begin{pmatrix} 30 \\ -20 \\ 40 \end{pmatrix}$$

$$= \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix}$$

$$\vec{C}_{new} = \vec{C}_{orig} - \frac{\vec{C}_{orig} \circ \vec{A}_{new}}{\vec{A}_{new} \circ \vec{A}_{new}} \vec{A}_{new} - \frac{\vec{C}_{orig} \circ \vec{B}_{new}}{\vec{B}_{new} \circ \vec{B}_{new}} \vec{B}_{new}$$

$$= \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} - \frac{\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \circ \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}}{\begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} \circ \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} - \frac{\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \circ \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix}}{\frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix} \circ \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix}} \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} - \frac{\begin{pmatrix} 2 & 3 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}}{\begin{pmatrix} 3 & -2 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} - \frac{\begin{pmatrix} 2 & 3 & -1 \end{pmatrix} \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix}}{\frac{1}{29} \begin{pmatrix} 30 & 9 & -18 \end{pmatrix} \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix}} \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} - \frac{-4}{29} \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} - \frac{\frac{105}{29}}{\frac{45}{29}} \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + \frac{1}{29} \begin{pmatrix} 12 \\ -8 \\ 16 \end{pmatrix} + \frac{1}{29} \begin{pmatrix} -70 \\ -21 \\ 42 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + \frac{1}{29} \begin{pmatrix} -58 \\ -29 \\ 58 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}.$$

Thus

$$\vec{A}_{new} = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} \qquad \vec{B}_{new} = \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix} \qquad \vec{C}_{new} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

and you can verify that these three vectors are orthogonal by taking the dot products between pairs:

$$\vec{A}_{new} \circ \vec{B}_{new} = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} \circ \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix} = \begin{pmatrix} 3 & -2 & 4 \end{pmatrix} \frac{1}{29} \begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix} = \frac{1}{29} [(3)(30) + (-2)(9) + (4)(-18)] = 0$$

$$\vec{A}_{new} \circ \vec{C}_{new} = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} \circ \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 & -2 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = [(3)(0) + (-2)(2) + (4)(1)] = 0$$

$$\vec{B}_{new} \circ \vec{C}_{new} = \frac{1}{29}\begin{pmatrix} 30 \\ 9 \\ -18 \end{pmatrix} \circ \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{29}\begin{pmatrix} 30 & 9 & -18 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{29}[(30)(0)+(9)(2)+(-18)(1)] = 0.$$

The same procedure works for continuous functions. So given two functions such as $f(x)$ and $g(x)$, the Gram-Schmidt procedure can be used to construct two orthogonal functions ($f_{new}(x)$ and $g_{new}(x)$:

$$f_{new}(x) = f_{orig}(x)$$
$$g_{new}(x) = g_{orig}(x) - \frac{g_{orig}(x) \circ f_{new}(x)}{f_{new}(x) \circ f_{new}(x)} f_{new}(x)$$

in which the inner product for continuous functions is

$$g(x) \circ f(x) = \int_{-\infty}^{\infty} g^*(x) f(x) dx.$$

As an example, consider $f(x) = x$ and $g(x) = x^3$ over interval of $x = -2$ to $x = 2$:

$$f_{new}(x) = f_{orig}(x) = x$$
$$g_{new}(x) = g_{orig}(x) - \frac{g_{orig}(x) \circ f_{new}(x)}{f_{new}(x) \circ f_{new}(x)} f_{new}(x) = g_{orig}(x) \ - \ \frac{\int_{-\infty}^{\infty} g^*_{orig}(x) f_{new}(x) dx}{\int_{-\infty}^{\infty} f^*_{new}(x) f_{new}(x) dx} f_{new}(x).$$

So

$$\begin{aligned} g_{new}(x) &= x^3 - \frac{\int_{-2}^{2} (x^3)(x) dx}{\int_{-2}^{2} (x)(x) dx} x = x^3 - \frac{\int_{-2}^{2} x^4 dx}{\int_{-2}^{2} x^2 dx} x \\ &= x^3 - \frac{\frac{1}{5}x^5|_{-2}^{2}}{\frac{1}{3}x^3|_{-2}^{2}} x = x^3 - \frac{\frac{1}{5}[2^5 - (-2)^5]}{\frac{1}{3}[2^3 - (-2)^3]} x = x^3 - \frac{\frac{64}{5}}{\frac{16}{3}} x \\ &= x^3 - \frac{12}{5} x. \end{aligned}$$

Thus $f_{new}(x) = x$ and $g_{new}(x) = x^3 - \frac{12}{5}x$.

To verify that $f_{new}(x)$ and $g_{new}(x)$ are orthogonal on the interval $x = -2$ to $x = 2$, check that the inner product between these two functions is zero:

$$\begin{aligned} g_{new}(x) \circ f_{new}(x) &= \int_{-2}^{2} \left(x^3 - \frac{12}{5}x \right)^* (x)dx = \int_{-2}^{2} \left(x^4 - \frac{12}{5}x^2 \right) dx \\ &= \frac{1}{5}x^5|_{-2}^{2} - \left(\frac{12}{5} \right) \frac{1}{3}x^3|_{-2}^{2} = \frac{64}{5} - \left(\frac{12}{5} \right) \frac{16}{3} = 0 \end{aligned}$$

as expected.

References

Weisstein, Eric W. "Gram-Schmidt Orthonormalization." From MathWorld–
A Wolfram Web Resource. http://mathworld.wolfram.com/Gram-SchmidtOrthonormalization.html

HMC Online Tutorial "The Gram-Schmidt Algorithm." From Harvey
Mudd College Mathematics. https://www.math.hmc.edu/calculus/tutorials/gramschmidt/