The Gram-Schmidt Orthonormalization Process

Def. Let V be an inner product space. A subset of V is an **orthonormal basis** for V if it is an ordered basis that is orthonormal.

Ex. The standard basis in $\mathbb{R}^n$ is an orthonormal basis.

Ex. The set $\left\{< \frac{3}{5}, \frac{4}{5} >, < \frac{4}{5}, -\frac{3}{5} >\right\}$ is an orthonormal basis for $\mathbb{R}^2$ since

$$\left\| < \frac{3}{5}, \frac{4}{5} > \right\| = \sqrt{\left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2} = 1, \quad \left\| < \frac{4}{5}, -\frac{3}{5} > \right\| = \sqrt{\left(\frac{4}{5}\right)^2 + \left(\frac{-3}{5}\right)^2} = 1$$

and $\left(< \frac{3}{5}, \frac{4}{5} >\right) \cdot \left(< \frac{4}{5}, -\frac{3}{5} >\right) = \frac{12}{25} - \frac{12}{25} = 0.$

Writing vectors in terms of an orthonormal basis can greatly simplify calculations.

Ex. Suppose $\{v_1, v_2, v_3, v_4\}$ is an orthonormal basis for an inner product space V and $v = 2v_1 - 3v_2 + v_3 + 4v_4$. Find $\|v\|$.

$$\|v\| = \sqrt{< v,\ v >} = \sqrt{< (2v_1 - 3v_2 + v_3 + 4v_4),\ (2v_1 - 3v_2 + v_3 + 4v_4) >}$$

But since $\{v_1, v_2, v_3, v_4\}$ is orthonormal we have that $< v_i,\ v_j > = \delta_{ij}$.

Thus we have:

$$\|v\| = \sqrt{2^2 + (-3)^2 + 1^2 + 4^2} = \sqrt{30}.$$

Theorem: Let V be an inner product space and $S = \{v_1, \dots, v_k\}$ be an orthogonal subset of nonzero vectors. If $v \in span(S)$ then $v = \sum_{i=1}^{k} \frac{<v, v_i>}{\|v_i\|^2} v_i$. In particular, if S is an orthonormal set then $v = \sum_{i=1}^{k} < v,\ v_i > v_i$.

The set $\{< v,\ v_i >\}$ are called the **Fourier coefficients** of v.

Proof: $v \in span(S) \implies \qquad v = \sum_{i=1}^{k} a_i v_i\ , \quad a_1, \dots, a_k \in \mathbb{R}.$

$$< v,\ v_j > = < \sum_{i=1}^{k} a_i v_i,\ v_j >$$

$$= < a_j v_j,\ v_j >$$

$$= a_j \|v_j\|^2.$$

$$\implies \qquad \frac{<v, v_j>}{\|v_j\|^2} = a_j.$$

Corollary: Let V be an inner product space and $S = \{v_1, \dots, v_k\}$ be an orthogonal subset of nonzero vectors. Then S is linearly independent.

Proof: Suppose that $a_1 v_1 + \cdots + a_k v_k = 0$. Let's show that $a_1 = \cdots = a_k = 0$.

By the previous theorem

$$a_j = \frac{<a_1 v_1 + \cdots + a_k v_k,\ v_j>}{\|v_j\|^2} = 0; \quad 1 \le j \le k.$$

Thus $S = \{v_1, \dots, v_k\}$ is linearly independent.

Ex. Let $v =< 3, -1, 2 >$ be written in the standard basis for $\mathbb{R}^3$. Write v in terms of the orthonormal basis

$\{w_1, w_2, w_3\} = \{\frac{1}{\sqrt{2}} < 1, 1, 0 >, \ \frac{1}{\sqrt{3}} < 1, -1, 1 >, \frac{1}{\sqrt{6}} < -1, 1, 2 >\}$. That is, find the Fourier coefficients of $v =< 3, -1, 2 >$ with respect to $\{w_1, w_2, w_3\}$.

$$a_1 =< v,\ w_1 > =<< 3, -1, 2 >,\ \ \frac{1}{\sqrt{2}} < 1, 1, 0 >>$$

$$= \frac{1}{\sqrt{2}}(3 - 1 + 0) = \frac{2}{\sqrt{2}} = \sqrt{2}.$$

$$a_2 =< v,\ w_2 > =<< 3, -1, 2 >,\ \ \frac{1}{\sqrt{3}} < 1, -1, 1 >>$$

$$= \frac{1}{\sqrt{3}}(3 + 1 + +2) = \frac{6}{\sqrt{3}} = 2\sqrt{3}.$$

$$a_3 =< v,\ w_3 > =<< 3, -1, 2 >,\ \ \frac{1}{\sqrt{6}} < -1, 1, 2 >>$$

$$= \frac{1}{\sqrt{6}}(-3 - 1 + 4) = 0.$$

Thus we have: $\quad v = \sqrt{2}w_1 + 2\sqrt{3}w_2.$

To check this result we see that:

$$\sqrt{2}w_1 + 2\sqrt{3}w_2 = \sqrt{2}(\frac{1}{\sqrt{2}} < 1, 1, 0 >) + 2\sqrt{3}(\frac{1}{\sqrt{3}} < 1, -1, 1 >)$$

$$=< 1, 1, 0 > +< 2, -2, 2 >$$

$$=< 3, -1, 2 >= v.$$

Our next goal is to show that given a basis for a finite dimensional inner product space V, we can find an orthonormal basis for V. We'll start by showing how we can create an orthogonal basis. Once we have an orthogonal basis we can create an orthonormal basis by dividing each basis vector by its length.

Let's see how we can create an orthogonal basis in $\mathbb{R}^2$ from a given basis $\{w_1, w_2\}$.

We start by letting $v_1 = w_1$.

Next we let $v_2 = w_2 - cw_1, \quad c \in \mathbb{R}$.

Now solve for c so that v_2 is orthogonal to w_1 (and hence orthogonal to v_1).

$$0 =< v_2,\ w_1 >=< w_2 - cw_1,\ w_1 >$$

$$=< w_2,\ w_1 > -c < w_1,\ w_1 >.$$

$$\Longrightarrow \qquad c < w_1,\ w_1 > =< w_2,\ w_1 >$$

$$c = \frac{<w_2, w_1>}{<w_1, w_1>}.$$

So $\qquad v_2 = w_2 - \frac{<w_2, w_1>}{<w_1, w_1>} w_1$ is orthogonal to $v_1 = w_1$.

This process can be generalized to create an orthogonal basis for any finite dimensional inner product space.

The process to create an orthonormal set in the next theorem is called the **Gram-Schmidt process**.

Theorem: Let V be an inner product space and $S = \{w_1, \dots, w_n\}$ a linearly independent subset of V. Define $S' = \{v_1, \dots, v_n\}$ where $v_1 = w_1$ and

$$v_k = w_k - \sum_{j=1}^{k-1} \frac{<w_k, v_j>}{\|v_j\|^2} v_j; \quad \text{for } 2 \le k \le n.$$

Then S' is an orthogonal set of nonzero vectors such that $span(S') = span(S)$.

Proof: We prove this theorem by mathematical induction on n, the number of vectors in S.

If $n = 1$ then $v_1 = w_1 \neq 0$, and $S' = \{v_1\}$ which is an orthogonal set.

Now assume the theorem is true for $S'_{k-1} = \{v_1, \dots, v_{k-1}\}$ and let's show that $S'_k = \{v_1, \dots, v_k\}$ is orthogonal where

$$(*) \qquad v_k = w_k - \sum_{j=1}^{k-1} \frac{<w_k, v_j>}{\|v_j\|^2} v_j.$$

If $v_k = 0$ then $w_k \in span(S'_{k-1}) = span(S_{k-1})$, which contradicts the assumption that $\{w_1, \dots, w_n\}$ is linearly independent. Thus $v_k \neq 0$.

For $1 \le i \le k-1$ we have:

$$< v_k, \ v_i > = < w_k, \ v_i > - \sum_{j=1}^{k-1} \frac{<w_k, v_j>}{\|v_j\|^2} < v_j, \ v_i >$$

$$= < w_k, \ v_i > - \frac{<w_k, \ v_i>}{\|v_i\|^2} < v_i, \ v_i >$$

$$= < w_k, \ v_i > - < w_k, \ v_i > = 0$$

since $< v_j, \ v_i > = 0, \ i \neq j$ by the induction hypothesis. Hence S'_k is orthogonal.

By $(*)$ $span(S'_k) \subseteq span(S_k)$.

But S'_k and S_k are both linearly independent sets so $\dim(span(S'_k)) = \dim(span(S_k)) \quad \Rightarrow \quad span(S'_k) = span(S_k)$.

Ex. Let $w_1 =< 4, 2, 2, 1 >$, $w_2 =< 2, 0, 0, 2 >$ and $w_3 =< 1, 1, -1, 1 >$ be vectors in $\mathbb{R}^4$. By a straight forward calculation one can check that $\{w_1, w_2, w_3\}$ is linearly independent. Use the Gram-Schmidt process to create an orthogonal set $\{v_1, v_2, v_3\}$ with the same span as $\{w_1, w_2, w_3\}$. Then normalize $\{v_1, v_2, v_3\}$ (ie divide by their lengths to create unit vectors) to create an orthonormal set $\{u_1, u_2, u_3\}$ with the same span as $\{w_1, w_2, w_3\}$. Find the Fourier coefficients of $v =< -1, 3, 1, -3 >$ with respect to the $\{u_1, u_2, u_3\}$.

$v_1 = w_1 =< 4, 2, 2, 1 >$.

$$v_2 = w_2 - \frac{< w_2,\ v_1 >}{\|v_1\|^2} v_1$$

$$=< 2, 0, 0, 2 > - \frac{<<2,0,0,2>,<4,2,2,1>>}{4^2+2^2+2^2+1^2} < 4, 2, 2, 1 >$$

$$=< 2, 0, 0, 2 > - \frac{2}{5} < 4, 2, 2, 1 >= \frac{2}{5} < 1, -2, -2, 4 >.$$

$$v_3 = w_3 - \frac{< w_3,\ v_1 >}{\|v_1\|^2} v_1 - \frac{< w_3,\ v_2 >}{\|v_2\|^2} v_2$$

$$=< 1, 1, -1, 1 > - \frac{<<1,1,-1,1>,<4,2,2,1>>}{4^2+2^2+2^2+1^2} < 4, 2, 2, 1 >$$

$$- \frac{<<1,1,-1,1>,\frac{2}{5}<1,-2,-2,4>>}{4} \frac{2}{5} < 1, -2, -2, 4 >$$

$$=< 1, 1, -1, 1 > - \frac{1}{5} < 4, 2, 2, 1 > - \frac{1}{5} < 1, -2, -2, 4 >$$

$$=< 0, 1, -1, 0 >.$$

So $v_1 =< 4, 2, 2, 1 >$

$v_2 = \frac{2}{5} < 1, -2, -2, 4 >$

$v_3 =< 0, 1, -1, 0 >$.

Normalizing these vectors we get:

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{5} < 4, 2, 2, 1 >$$

$$u_2 = \frac{v_2}{\|v_2\|} = \frac{1}{5} < 1, -2, -2, 4 >$$

$$u_3 = \frac{v_3}{\|v_3\|} = \frac{1}{\sqrt{2}} < 0, 1, -1, 0 >.$$

To find the Fourier coefficients:

$$< v,\ u_1 > = << -1, 3, 1, -3 >, \frac{1}{5} < 4, 2, 2, 1 >>$$

$$= \frac{1}{5}(-4 + 6 + 2 - 3) = \frac{1}{5}$$

$$< v,\ u_2 > = << -1, 3, 1, -3 >, \frac{1}{5} < 1, -2, -2, 4 >>$$

$$= \frac{1}{5}(-1 - 6 - 2 - 12) = -\frac{21}{5}$$

$$< v,\ u_3 > = << -1, 3, 1, -3 >, \frac{1}{\sqrt{2}} < 0, 1, -1, 0 >>$$

$$= \frac{1}{\sqrt{2}}(0 + 3 - 1 + 0) = \frac{2}{\sqrt{2}} = \sqrt{2}\ .$$

Notice that since $v =< -1, 3, 1, -3 >$ is not in the $span\{u_1, u_2, u_3\}$

$$v \neq a_1 u_1 + a_2 u_2 + a_3 u_3$$

for any $a_1, a_2, a_3 \in \mathbb{R}$. Thus the best we can say is

$$v =< -1, 3, 1, -3 > \approx \frac{1}{5}(\frac{1}{5} < 4, 2, 2, 1 >) - \frac{21}{5}\left(\frac{1}{5} < 1, -2, -2, 4 >\right)$$

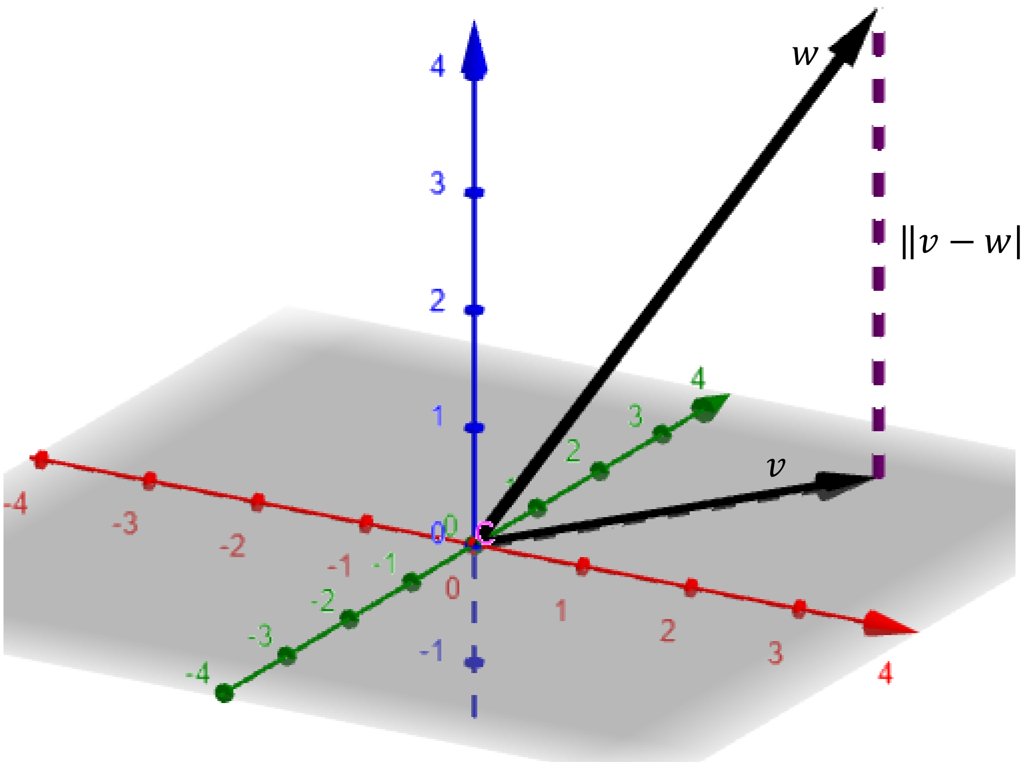
$$+\sqrt{2}(\frac{1}{\sqrt{2}} < 0, 1, -1, 0 >).$$

In fact, it's the "best" possible approximation of $< -1, 3, 1, -3 >$.

Suppose we measure the distance between two vectors $v, w \in V$ by

$$dist(v, w) = \|v - w\|.$$

We can then ask the question, suppose $w \in V$ and $v \in V' \subseteq V$, what vector $v \in V' \subseteq V$ minimizes $dist(v, w) = \|v - w\|$?



As we will see in the next theorem, if $\{v_1, v_2, \dots, v_k\}$ is an orthonormal basis for $V' \subseteq V$ then the vector $v \in V' \subseteq V$ that minimizes $dist(v, w) = \|v - w\|$ for $w \in V$, is the vector $v = \alpha_1 v_1 + \cdots + \alpha_k v_k$, where α_i is the i^{th} Fourier coefficient, $\alpha_i =< w, v_i >$ of w. Thus $v =< w, v_1 > v_1 + \cdots +< w, v_k > v_k$ is the best approximation of w in the sense that $dist(v, w) = \|v - w\|$ is as small as it can be among vectors in $V' \subseteq V$.

Theorem: Let V be a finite dimensional inner product space and $V' \subseteq V$ a subspace of V. Suppose that $\{v_1, \dots, v_n\}$ is an orthonormal basis for V and $\{v_1, \dots, v_k\}$, where $k \leq n$, is an orthonormal basis for $V' \subseteq V$. Given any vector $w \in V$ the vector in $v \in V'$ that minimizes $dist(v, w) = \|v - w\|$ is

$$v = < w, v_1 > v_1 + \dots + < w, v_k > v_k.$$

Proof. Since $\{v_1, \dots, v_n\}$ is an orthonormal basis for V we can write:

$$w = < w, v_1 > v_1 + \dots + < w, v_n > v_n.$$

Since $\{v_1, \dots, v_k\}$ is a basis for $V' \subseteq V$ we can represent any vector $v \in V'$ by

$$v = \alpha_1 v_1 + \dots + \alpha_k v_k.$$

To find a vector that minimizes $dist(v, w) = \|v - w\|$ it is equivalent to find one that minimizes $\left(dist(v, w)\right)^2 = \|v - w\|^2$.

$$\|v - w\|^2 = < (v - w), (v - w) >$$

$$= < (\alpha_1 - < w, v_1 >)v_1 + \dots + (\alpha_k - < w, v_k >)v_k - \dots - < w, v_n > v_n,$$

$$(\alpha_1 - < w, v_1 >)v_1 + \dots + (\alpha_k - < w, v_k >)v_k - \dots - < w, v_n > v_n >.$$

Since $< v_i, v_j > = \delta_{ij}$ we get:

$$\|v - w\|^2 = (\alpha_1 - < w, v_1 >)^2 + \dots + (\alpha_k - < w, v_k >)^2 + \dots + < w, v_n >^2.$$

The RHS has a minimum value at:

$\alpha_1 - < w, v_1 > = 0, \dots ,\ \alpha_k - < w, v_k > = 0$ or

$\alpha_1 = < w, v_1 >, \dots, \alpha_k = < w, v_k >$.

Hence: $v = < w, v_1 > v_1 + \dots + < w, v_k > v_k$ minimizes $dist(v, w) = \|v - w\|$.

Ex. Let $V = \{polynomials\ with\ real\ coefficients\ on\ [0.1]\}$ with the inner product $< f,\ g >= \int_0^1 f(x)g(x)dx$. Consider the subspace $W \subseteq V$ where $W = \{polynomials\ with\ real\ coefficients\ of\ degree\ \leq 2\ on\ [0,1]\}$ with the ordered basis $\{1, 2x, 3x^2\}$. Find an orthogonal basis for W and then normalize it to obtain an orthonormal basis for W.

$$w_1 = 1, \quad w_2 = 2x, \quad w_3 = 3x^2.$$

$$v_1 = w_1 = 1.$$

$$v_2 = w_2 - \frac{<w_2, v_1>}{\|v_1\|^2} v_1$$

$$= 2x - \frac{\int_0^1 2x(1)dx}{\int_0^1 1^2 dx}(1)$$

$$\int_0^1 2x(1)dx = x^2\Big|_0^1 = 1; \qquad \int_0^1 1^2 dx = x\Big|_0^1 = 1 \qquad \text{so}$$

$$v_2 = 2x - \frac{1}{1}(1) = 2x - 1.$$

$$v_3 = w_3 - \frac{< w_3,\ v_1 >}{\|v_1\|^2} v_1 - \frac{< w_3,\ v_2 >}{\|v_2\|^2} v_2$$

$$= 3x^2 - \frac{\int_0^1 3x^2(1)dx}{\int_0^1 1^2 dx}(1) - \frac{\int_0^1 3x^2(2x-1)dx}{\int_0^1 (2x-1)^2 dx}(2x-1)$$

$$= 3x^2 - \frac{x^3\Big|_0^1}{1}(1) - \frac{\int_0^1 (6x^3 - 3x^2)dx}{\int_0^1 (4x^2 - 4x + 1)dx}(2x-1)$$

$$= 3x^2 - 1 - \frac{\frac{1}{2}}{\frac{1}{3}}(2x-1) = 3x^2 - 3x + \frac{1}{2}.$$

So $\{1,\ (2x-1),\ (3x^2-3x+\frac{1}{2})\}$ is an orthogonal basis for W.

Now we normalize $\{v_1, v_2, v_3\}$ to get $\{u_1, u_2, u_3\}$ an orthonormal basis for W.

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{\int_0^1 1^2 dx}} = \frac{1}{1} = 1$$

$$u_2 = \frac{v_2}{\|v_2\|} = \frac{2x-1}{\sqrt{\int_0^1 (2x-1)^2 dx}} = \frac{(2x-1)}{\sqrt{\frac{1}{3}}} = \sqrt{3}(2x-1)$$

$$u_3 = \frac{v_3}{\|v_3\|} = \frac{3x^2-3x+\frac{1}{2}}{\sqrt{\int_0^1 \left(3x^2-3x+\frac{1}{2}\right)^2 dx}}$$

$$= \frac{3x^2-3x+\frac{1}{2}}{\sqrt{\int_0^1 (9x^4-18x^3+12x^2-3x+\frac{1}{4}) dx}}$$

$$= \frac{3x^2-3x+\frac{1}{2}}{\sqrt{\frac{1}{20}}} = \sqrt{5}(6x^2-6x+1).$$

Ex. Find the Fourier coefficients of $f(x) = 2 - 3x^2$ using the orthonormal basis found in the previous example and write $f(x)$ as a linear combination of this orthonormal basis.

$u_1 = 1$, $u_2 = \sqrt{3}(2x - 1)$, $u_3 = \sqrt{5}(6x^2 - 6x + 1)$.

$$< f(x),\, u_1 >= \int_0^1 (2 - 3x^2)(1)dx = (2x - x^3)\Big|_0^1 = 1\,.$$

$$< f(x),\, u_2 >= \int_0^1 (2 - 3x^2)\big(\sqrt{3}(2x - 1)\big)dx$$

$$= \sqrt{3}\ \int_0^1 (-6x^3 + 3x^2 + 4x - 2)dx$$

$$= \sqrt{3}\ \left(-\frac{3x^4}{2} + x^3 + 2x^2 - 2x\right)\Big|_0^1$$

$$= -\frac{\sqrt{3}}{2}$$

$$< f(x),\, u_3 >= \int_0^1 (2 - 3x^2)\left(\sqrt{5}(6x^2 - 6x + 1)\right)dx$$

$$= \sqrt{5}\ \int_0^1 (-18x^4 + 18x^3 + 9x^2 - 12x + 2)dx$$

$$= \sqrt{5}\ \left(-\frac{18}{5}x^5 + \frac{9}{2}x^4 + 3x^3 - 6x^2 + 2x\right)\Big|_0^1$$

$$= -\frac{\sqrt{5}}{10}\,.$$

So we can express $f(x)$ as a linear combination of $\{u_1, u_2, u_3\}$ as:

$$f(x) = 2 - 3x^2 = 1(1) - \frac{\sqrt{3}}{2}\left(\sqrt{3}(2x - 1)\right) - \frac{\sqrt{5}}{10}\left(\sqrt{5}(6x^2 - 6x + 1)\right).$$