Definition: Let V be an inner product space. We say $\mathbf{v}, \mathbf{w} \in V$ are **orthogonal** if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. If $S \subseteq V$ and for all $\mathbf{v}, \mathbf{w} \in S$ and $\mathbf{v} \neq \mathbf{w}$ we have $\langle \mathbf{v}, \mathbf{w} \rangle = 0$, then we say S is an **orthogonal** set. If $\mathbf{u} \in V$ and $\|\mathbf{u}\| = 1$, we say $\mathbf{u}$ is a **unit vector**. Notice that any non-zero vector $\mathbf{v} \in V$ can be **normalized** (i.e. $\mathbf{u} = \mathbf{v}/\|\mathbf{v}\|$) and turned into a unit vector. If S is an orthogonal set of unit vectors, S is called an **orthonormal** set. It is not hard to show that an orthogonal set is linearly independent if and only if the zero vector does not belong to the set. In particular, an orthonormal set is always linearly independent.

Coordinates and norms work particularly well in the context of orthogonal or orthonormal bases. Suppose that $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an orthogonal basis. Let $\mathbf{v} \in V$ and write $\mathbf{v} = \sum\limits_{i=1}^{n} c_i \mathbf{v}_i$. Then we have that $\langle \mathbf{v}, \mathbf{v}_i \rangle = \left\langle \sum\limits_{j=1}^{n} c_j \mathbf{v}_j, \mathbf{v}_i \right\rangle = \sum\limits_{j=1}^{n} c_j \langle \mathbf{v}_j, \mathbf{v}_i \rangle = c_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle = c_i \|\mathbf{v}_i\|^2$ where the sum collapses out since the inner products for $i \neq j$ yield zeros. So this means that $\mathbf{v} = \sum\limits_{i=1}^{n} \dfrac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i$. So coordinates can be computed just using inner products! When our basis β is orthonormal (i.e. $\|\mathbf{v}_i\| = 1$) this is even nicer: $\mathbf{v} = \sum\limits_{i=1}^{n} \langle \mathbf{v}, \mathbf{v}_i \rangle \mathbf{v}_i$. In this case the *coordinates are the inner products with* $\mathbf{v}$.

Suppose that $S = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ is orthogonal. Then

$$\|\mathbf{w}_1 + \cdots + \mathbf{w}_n\|^2 = \left\langle \sum_{i=1}^{n} \mathbf{w}_i, \sum_{j=1}^{n} \mathbf{w}_j \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^{n} \langle \mathbf{w}_i, \mathbf{w}_i \rangle = \|\mathbf{w}_1\|^2 + \cdots + \|\mathbf{w}_n\|^2$$

where one of the sums drops out because $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = 0$ if $i \neq j$. This then implies that if $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an orthonormal basis, then $\|\mathbf{v}\|^2 = \sum_{i=1}^{n} \|\langle \mathbf{v}, \mathbf{v}_i \rangle \mathbf{v}_i\|^2 = \sum\limits_{i=1}^{n} |\langle \mathbf{v}, \mathbf{v}_i \rangle|^2$ (i.e. the length of $\mathbf{v}$ is the sum of the squares of the lengths of its coordinates.

Let $\alpha = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ be a linearly independent set. Recursively define a set $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ as follows:

$$\mathbf{v}_1 = \mathbf{w}_1 \qquad\qquad \mathbf{v}_{n+1} = \mathbf{w}_{n+1} - \mathrm{proj}_{\mathbf{v}_1}(\mathbf{w}_{n+1}) - \cdots - -\mathrm{proj}_{\mathbf{v}_n}(\mathbf{w}_{n+1})$$

where $\mathrm{proj}_{\mathbf{v}}(\mathbf{w}) = \dfrac{\langle \mathbf{w}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2} \mathbf{v}$ is the orthogonal projection of $\mathbf{w}$ onto $\mathbf{v}$. These sets then have the property that $\mathrm{span}\{\mathbf{w}_1, \dots, \mathbf{w}_j\} = \mathrm{span}\{\mathbf{v}_1, \dots, \mathbf{v}_j\}$ for each $j = 1, \dots, n$ and $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is orthogonal (and linearly independent). In particular, if α is a basis for V, then β is an orthogonal basis for V. Computing the $\mathbf{v}_i$'s is called the **Gram-Schmidt** orthogonalization process. If we add in the additional step of normalizing the elements of β, then we will have an orthonormal basis for V.

One consequence of the Gram-Schmidt algorithm is that *every finite dimensional inner product space has an orthonormal basis.* Let $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be an orthonormal basis for V. Recall that $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ with 1 in the i-th position. Define a linear map φ that sends $\mathbf{v}_i$ to $\mathbf{e}_i$ for $i = 1, \dots, n$. This map is an isomorphism (since we send a basis to a basis) and $\varphi(\langle \mathbf{v}, \mathbf{w} \rangle) = \varphi(\mathbf{v}) \bullet \varphi(\mathbf{w})$ (it preserves inner products). Thus $V \cong \mathbb{F}^n$ (isomorphic not only as vector spaces but also as inner product spaces). So in a sense, the only inner product one can have on a finite dimensional inner product space is the dot product (up to isomorphism).

It is worth mentioning that if you start with merely a spanning set, Gram-Schmidt will still produce an orthogonal basis – if we toss out the zero vector anytime it appears (if a vector is linearly dependent on previous vectors, once we subtract out projections, nothing will be left).

Example: Consider $S = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4\} = \{(1,0,0,i), (1+i,1,1,1+i), (3+i,1,1,1+3i), (0,1+2i,0,1)\}$.

- $\mathbf{v}_1 = \mathbf{w}_1 = (1,0,0,i)$
- $\mathbf{v}_2 = \mathbf{w}_2 - \dfrac{\mathbf{w}_2 \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = (1+i,1,1,1+i) - \dfrac{2}{2}(1,0,0,i) = (i,1,1,1)$
- Notice that $\mathbf{w}_3 - \dfrac{\mathbf{w}_3 \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\mathbf{w}_3 \bullet \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 = (3+i,1,1,1+3i) - \dfrac{6}{2}(1,0,0,i) - \dfrac{4}{4}(i,1,1,1) = (0,0,0,0)$.
  This means that $\mathbf{w}_3 \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \text{span}\{\mathbf{w}_1, \mathbf{w}_2\}$. So we can toss out $\mathbf{w}_3$ (and $\mathbf{0}$).
- $\mathbf{v}_3 = \mathbf{w}_4 - \dfrac{\mathbf{w}_4 \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\mathbf{w}_4 \bullet \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 = (0,1+2i,0,1) - \dfrac{-i}{2}(1,0,0,i) - \dfrac{2+2i}{4}(i,1,1,1) = \dfrac{1}{2}(1,1+3i,-1-i,-i)$

So $\text{span}(S) = \text{span}\beta$ where $\beta = \left\{(1,0,0,i), (i,1,1,1), \dfrac{1}{2}(1,1+3i,-1-i,-i)\right\}$ is an orthogonal basis. We could also normalize the vectors in β and get $\alpha = \left\{\dfrac{1}{\sqrt{2}}(1,0,0,i), \dfrac{1}{2}(i,1,1,1), \dfrac{1}{\sqrt{14}}(1,1+3i,-1-i,-i)\right\}$. α is then an orthonormal basis for $\text{span}(S)$.

Example: Consider the inner product $\langle f, g\rangle = \displaystyle\int_{-1}^{1} f(x)g(x)\,dx$ defined on $\mathbb{R}[x]$. If we run the Gram-Schmidt algorithm (unnormalized) on $\{1, x, x^2, \dots\}$ and then rescale them so that $\|P_n(x)\|^2 = \dfrac{2}{2n+1}$, we will get a collection of polynomials known as *Legendre polynomials* (we will denote them $P_0(x), P_1(x)$, etc.). In particular...

- $f_0(x) = 1$. We then need $2(?)^2 = (?)^2 \int_{-1}^{1} 1^2\,dx = (?)^2\|f_0(x)\|^2 = \dfrac{2}{2(0)+1}$, so $? = 1$ and thus $P_0(x) = f_0(x) = 1$.
- $f_1(x) = x - \dfrac{\langle x, 1\rangle}{\|1\|^2}1 = x - \dfrac{\int_{-1}^{1} x\cdot 1\,dx}{\int_{-1}^{1} 1^2\,dx}1 = x - \dfrac{0}{2}1 = x$. We then need $\dfrac{2}{3}(?)^2 = \|(?)^2 f_1(x)\|^2 = \dfrac{2}{2(1)+1}$ so $? = 1$ and thus $P_1(x) = f_1(x) = x$.
- $f_2(x) = x^2 - \dfrac{\langle x^2, 1\rangle}{\|1\|^2}1 - \dfrac{\langle x^2, x\rangle}{\|x\|^2}x = x^2 - \dfrac{\int_{-1}^{1} x^2\cdot 1\,dx}{\int_{-1}^{1} 1^2\,dx}1 - \dfrac{\int_{-1}^{1} x^2\cdot x\,dx}{\int_{-1}^{1} x^2\,dx}x = x^2 - \dfrac{2/3}{2}1 - \dfrac{0}{2/3}x = x^2 - \dfrac{1}{3}$.
  We then need $\dfrac{8}{45}(?)^2 = (?)^2\int_{-1}^{1}(x^2-1/3)^2\,dx = \|?\,f_2(x)\|^2 = \dfrac{2}{2(2)+1} = \dfrac{2}{5}$ so $(?)^2 = 9/4$ and thus $? = 3/2$. Thus $P_2(x) = (3/2)f_2(x) = \dfrac{3}{2}x^2 - \dfrac{1}{2}$.