4.2 The Gram-Schmidt Process

So, we've learned orthogonal and orthonormal bases are important. But how do we construct them?

Gram-Schmidt Process

Let W denote a finite dimensional inner product space. Assume some basis $\vec{w}_1, \ldots, \vec{w}_n$ of W, where $n = \dim W$.

Our goal is to construct an orthogonal basis $\vec{v}_1, \ldots, \vec{v}_n$. So, set $\vec{v}_1 := \vec{w}_1$. Note that $\vec{v}_1 \neq \vec{0}$.

Next, working with $\vec{w}_2$, and insisting $\langle \vec{v}_2, \vec{v}_1 \rangle = 0$, we arrange this by subtracting from $\vec{w}_2$ a suitable multiple of $\vec{v}_1$:

$\vec{v}_2 = \vec{w}_2 - c\vec{v}_1$, where c is yet to be determined.

$0 = \langle \vec{v}_2, \vec{v}_1 \rangle = \langle \vec{w}_2 - c\vec{v}_1, \vec{v}_1 \rangle = \langle \vec{w}_2, \vec{v}_1 \rangle - c\langle \vec{v}_1, \vec{v}_1 \rangle = \langle \vec{w}_2, \vec{v}_1 \rangle - c\left|\vec{v}_1\right|^2$, requiring $c = \frac{\langle \vec{w}_2, \vec{v}_1 \rangle}{|\vec{v}_1|^2}$.

Therefore: $\vec{v}_2 := \vec{w}_2 - \frac{\langle \vec{w}_2, \vec{v}_1 \rangle}{|\vec{v}_1|^2}\vec{v}_1$.

How does the linearity of $\vec{v}_1 = \vec{w}_1$ and $\vec{w}_2$ ensure that $\vec{v}_2 \neq \vec{0}$?

Similarly, we would find $\vec{v}_3 := \vec{w}_3 - \frac{\langle \vec{w}_3, \vec{v}_1 \rangle}{|\vec{v}_1|^2}\vec{v}_1 - \frac{\langle \vec{w}_3, \vec{v}_2 \rangle}{|\vec{v}_2|^2}\vec{v}_2$.

More generally, suppose we've already constructed mutually orthogonal $\vec{v}_1, \ldots, \vec{v}_{k-1}$ as linear combinations of $\vec{w}_1, \ldots, \vec{w}_{k-1}$.

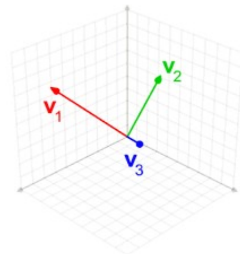
The next orthogonal basis element $\vec{v}_k$ will be obtained from $\vec{w}_k$ by subtracting off a suitable linear combination of the previous orthogonal basis elements:

$\vec{v}_k = \vec{w}_k - c_1\vec{v}_1 - \ldots - c_{k-1}\vec{v}_{k-1}$.

And since $\vec{v}_1, \ldots, \vec{v}_{k-1}$ are already orthogonal, for each $j = 1, \ldots, k-1$, we use the orthogonality constraint:

$0 = \langle \vec{v}_k, \vec{v}_j \rangle = \langle \vec{w}_k, \vec{v}_j \rangle - c_j\langle \vec{v}_j, \vec{v}_j \rangle$ requiring $c_j = \frac{\langle \vec{w}_k, \vec{v}_j \rangle}{|\vec{v}_j|^2}$.

In this fashion, we establish the general Gram-Schmidt formula: $\vec{v}_k := \vec{w}_k - \sum_{j=1}^{k-1} \frac{\langle \vec{w}_k, \vec{v}_j \rangle}{|\vec{v}_j|^2}\vec{v}_j, \quad k = 1, \ldots, n. \qquad (*)$



(See animation in class)

To form an orthonormal basis, given a basis $\vec{w}_1, \ldots, \vec{w}_n$, we replace each orthogonal basis vector in the Gram-Schmidt formula $(*)$ by its normalized version $\vec{u}_j = \frac{\vec{v}_j}{|\vec{v}_j|}$.

Example: Observe that $\vec{w}_1 = (1,1,-1)$, $\vec{w}_2 = (1,0,2)$, and $\vec{w}_3 = (2,-2,3)$ form a basis of $\mathbb{R}^3$ (feel free to check). Construct an orthogonal basis (with respect to the standard dot product) using the Gram-Schmidt process.

$$\vec{v}_1 := \vec{w}_1 = (1,1,-1).$$

$$\vec{v}_2 = \vec{w}_2 - \frac{\vec{w}_2 \cdot \vec{v}_1}{|\vec{v}_1|^2}\vec{v}_1 = (1,0,2) - \left(-\frac{1}{3}\right)(1,1,-1) = \left(\frac{4}{3}, \frac{1}{3}, \frac{5}{3}\right).$$

$$\vec{v}_3 = \vec{w}_3 - \frac{\vec{w}_3 \cdot \vec{v}_1}{|\vec{v}_1|^2}\vec{v}_1 - \frac{\vec{w}_3 \cdot \vec{v}_2}{|\vec{v}_2|^2}\vec{v}_2 = (2,-2,3) - \left(-\frac{3}{3}\right)(1,1,-1) - \left(\frac{7}{\frac{14}{3}}\right)\left(\frac{4}{3}, \frac{1}{3}, \frac{5}{3}\right) = \left(1, -\frac{3}{2}, -\frac{1}{2}\right).$$

What are the corresponding orthonormal basis vectors?

$$|\vec{v}_1| = \sqrt{3}, \qquad |\vec{v}_2| = \sqrt{\frac{14}{3}}, \qquad |\vec{v}_3| = \sqrt{\frac{7}{2}}.$$

Therefore: $\vec{u}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$, $\vec{u}_2 = \left(\frac{4}{\sqrt{42}}, \frac{1}{\sqrt{42}}, \frac{5}{\sqrt{42}}\right)$, $\vec{u}_3 = \left(\frac{2}{\sqrt{14}}, -\frac{3}{\sqrt{14}}, -\frac{1}{\sqrt{14}}\right)$.

According to a previous theorem, every finite dimensional vector space (except $\{\vec{0}\}$) admits a basis. Therefore, given the Gram-Schmidt process, we have the following...

Theorem: Every nonzero finite dimensional inner product space has an orthonormal basis.

Modifications of the Gram-Schmidt Process

Recall the general Gram-Schmidt: $\vec{v}_k := \vec{w}_k - \sum_{j=1}^{k-1} \frac{\langle \vec{w}_k, \vec{v}_j \rangle}{|\vec{v}_j|^2} \vec{v}_j, \quad k = 1, \ldots, n.$

If you substitute into this: $\vec{v}_j = |\vec{v}_j| \vec{u}_j$, and then solve for $\vec{w}_k$, then the original basis $\vec{w}_i$ can be expressed in terms of the orthonormal basis $\vec{u}_i$ via a triangular system.

$$
\begin{aligned}
\vec{w}_1 &= r_{11}\vec{u}_1, \\
\vec{w}_2 &= r_{12}\vec{u}_1 + r_{22}\vec{u}_2, \\
\vec{w}_3 &= r_{13}\vec{u}_1 + r_{23}\vec{u}_2 + r_{33}\vec{u}_3, \\
\vdots \quad & \quad \vdots \qquad\quad \vdots \qquad\quad \ddots \\
\vec{w}_n &= r_{1n}\vec{u}_1 + r_{2n}\vec{u}_2 + \ldots + r_{nn}\vec{u}_n,
\end{aligned}
$$

where $r_{ij} = \langle \vec{w}_j, \vec{u}_i \rangle$. $\qquad (**)$

For instance: $\vec{v}_1 := \vec{w}_1$, so substituting, we have $\vec{w}_1 = |\vec{v}_1| \vec{u}_1 = |\vec{w}_1| \vec{u}_1 = \frac{1}{|\vec{w}_1|} |\vec{w}_1|^2 \vec{u}_1$

$$= \frac{1}{|\vec{w}_1|} \langle \vec{w}_1, \vec{w}_1 \rangle \vec{u}_1 = \langle \vec{w}_1, \vec{u}_1 \rangle \vec{u}_1 = r_{11}\vec{u}_1.$$

The value for r_{ij} can be shown more generally by taking the inner product of each of these $\vec{w}_j$ equations with the orthonormal basis vectors $\vec{u}_i$ for $i \leq j$.

Taking into account the orthonormality constraints observe:

$\langle \vec{w}_j, \vec{u}_i \rangle = \langle r_{1j}\vec{u}_1 + \ldots + r_{jj}\vec{u}_j, \vec{u}_i \rangle = r_{1j}\langle \vec{u}_1, \vec{u}_i \rangle + \ldots + r_{jj}\langle \vec{u}_n, \vec{u}_i \rangle = r_{ij}$, and hence: $r_{ij} = \langle \vec{w}_j, \vec{u}_i \rangle$.

$\vec{v}_2 := \vec{w}_2 - \frac{\langle \vec{w}_2, \vec{v}_1 \rangle}{|\vec{v}_1|^2} \vec{v}_1$, so substituting: $\vec{v}_2 := \vec{w}_2 - \frac{\langle \vec{w}_2, |\vec{v}_1| \vec{u}_1 \rangle}{|\vec{v}_1|^2} |\vec{v}_1| \vec{u}_1 = \vec{w}_2 - \langle \vec{w}_2, \vec{u}_1 \rangle \vec{u}_1 = \vec{w}_2 - r_{12}\vec{u}_1$

Therefore: $\vec{w}_2 = r_{12}\vec{u}_1 + \vec{v}_2 = r_{12}\vec{u}_1 + |\vec{v}_2| \vec{u}_2$. So apparently: $|\vec{v}_2| = r_{22} = \langle \vec{w}_2, \vec{u}_2 \rangle$.

We will use the system $(**)$ in the next section to recharacterize the Gram-Schmidt process as matrix factorization.