7.2 Inner products and orthonormal bases

We are now pivoting to the idea of **orthogonality** and **inner products**, specifically in the application of vectors in $\mathbb{R}^n$. Let's cover some basics first that you're probably familiar with.

Inner products, orthogonality

- **Inner products**: some computation between two elements to yield a scalar quantity, denoted $< v, w >$ between elements v, w
  - a **dot product** is a specific type of inner product between two vectors, where given two $\mathbb{R}^n$ vectors, v, w, $< v, w >= v \cdot w = v^T w = v_1^T w_1 + v_2^T w_2 + ... + v_n^T w_n$, which results in a scalar quantity representing how those two $\mathbb{R}^n$ vectors exist in similar directions in $\mathbb{R}^n$
  - $cos(\theta) = \frac{<v,w>}{|v||w|}$, where θ is the angle between two vectors, v, w
  - inner products are very broad, for example you can take the inner product of two functions $f(x), g(x)$ using a definite integral.
  - in this class, we only focus on inner products as dot products, but it is important to note that they are not interchangable terms
- **Orthogonality**: when two vectors do not exist in the same direction in $\mathbb{R}^n$
  - Two vectors, v, w, are orthogonal when their dot product $= 0$: $< v, w >= v \cdot w = 0$
  - if two vectors are orthogonal, they must also be linearly independent, but NOT the other way around
    * orthogonality $\Rightarrow$ linear independence
    * linear independence $\not\Rightarrow$ orthogonality
- **Orthonormal**: two vectors v, w are orthonormal when they are 1. **orthogonal** and 2. **normal**, meaning they have length$= 1$
  - vectors are **normalized** when you solve for their **unit vector**, e_n, by dividing an entire vector by its magnitude
  - **unit vector calculation**: $e_n = \frac{v_n}{|v_n|} = \frac{1}{|v_n|}[v_1, v_2, ..., v_n]^T$ where $|v_n| = \sqrt{v_1^2 + v_2^2 + ... + v_n^2}$
- **Projections**: a projection of a vector can be onto another vector, or onto multiple vectors/a vector space. Projections result in the computation of how much some vector of interest v, exists in the direction of the other vectors/vector space
  - $proj_w v$ is the calculation of how much the vector, v, exists in the direction of the **vector** w
    * $proj_w v = < v, e_w > e_w = \frac{<v,w>}{<w,w>} w$, where e_w is the unit vector of w. Note that these two expressions are equivalent, and you can either
      solve for $proj_w v$ by explicitly solving for e_w first, computing the dot product of v and e_w, and then multiplying by e_w, or you can
      solve for the dot product of both v, w and w, w and multiply by the original w vector
  - $proj_H v$ is the calculation of how much the vector, v, exists in the direction of the **vector space** H, which has a basis $\{h_1, h_2, h_3\}$

  * $proj_H v = <v, e_{h,1}> e_{h,1} + <v, e_{h,2}> e_{h,2} + <v, e_{h,3}> e_{h,3}$
  $= \frac{<v,h_1>}{<h_1,h_1>} h_1 + \frac{<v,h_2>}{<h_2,h_2>} h_2 + \frac{<v,h_3>}{<h_3,h_3>} h_3$
- Note that all projection calculations will *always* yield a vector, not a scalar

Now that we've covered the foundations of inner products, orthogonality, and projections, what are we using them for? In this class, we want to be able to solve for **orthonormal bases** for a given vector space, a.k.a. a basis for some V that has orthonomal basis vectors. To do this, we will use the **Gram Schmidt process**. The best way to learn this process is to see a thorough example:

Gram Schmidt example

Solve for the orthonormal basis for the vector space:

$$V = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix} \right\}$$

Step 1: Denote vectors $v_1, v_2, ..., v_k$ and find e_1

Your choice of v_1, v_2, v_3 in this example is arbitrary, but we let's just say

$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \end{bmatrix}$, and $v_3 = \begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix}$, so

$$e_1 = \frac{v_1}{|v_1|} \rightarrow |v_1| = \sqrt{1^2 + 0^2 + 0^2 + 1^2} = \sqrt{2}$$

$$e_1 = \frac{v_1}{\sqrt{2}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = e_1$$

Step 2: Project v_2 onto e_1 and subtract to get w_2

Here, we are taking e_1 as our "reference", and now want to find how much of v_2 goes in the same direction as e_1, and we **want to remove the portion that goes in the same direction**. So, we will subtract $proj_{e_1} v_2$ from v_2 to get the portion of v_2 completely orthogonal to e_1, which we can denote as w_2

$$w_2 = v_2 - proj_{e_1} v_2 = v_2 - <v_2, e_1> e_1 =$$

$$\begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \end{bmatrix} - \left(2\left(\tfrac{1}{\sqrt{2}}\right) + 0(0) + 0(0) + -2\left(\left(\tfrac{1}{\sqrt{2}}\right)\right)\right) \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} =$$

$$\begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \end{bmatrix} - (0) \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \end{bmatrix} = w_2$$

In this step, w_2 was the exact same as v_2, because e_1 and v_2 were orthogonal from the start of the problem, so $proj_{e_1} v_2 = 0$. This is not always the case, obviously.

Step 3: Find e_2

$$e_2 = \frac{w_2}{|w_2|} \to |w_2| = \sqrt{2^2 + 0^2 + 0^2 + (-2)^2} = \sqrt{8}$$

$$e_2 = \frac{w_2}{\sqrt{8}} = \begin{bmatrix} \frac{2}{\sqrt{8}} \\ 0 \\ 0 \\ \frac{-2}{\sqrt{8}} \end{bmatrix} = e_2$$

Step 4: Project v_3 onto e_1 AND e_2 and subtract to get w_3

Recall that we are forming and orthonormal basis comprised of $\{e_1, e_2, e_3\}$, which means that all three vectors must be *mutually orthonormal*. From steps 2 and 3, we know that e_1 and e_2 are orthonormal to each other, but with e_3, we must subtract away the projection with both e_1 AND e_2.

$$w_3 = v_3 - proj_{e_1} v_3 - proj_{e_2} v_3 = v_3 - < v_3, e_1 > e_1 - < v_3, e_2 > e_2 =$$

$$\begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix} - (2(\tfrac{1}{\sqrt{2}}) + 1(0) + 1(0) + -1(\tfrac{1}{\sqrt{2}})) \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} - (2(\tfrac{2}{\sqrt{8}}) + 1(0) + 1(0) + -1(\tfrac{-2}{\sqrt{8}}) \begin{bmatrix} \frac{2}{\sqrt{8}} \\ 0 \\ 0 \\ \frac{-2}{\sqrt{8}} \end{bmatrix} =$$

$$\begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix} - (\tfrac{1}{\sqrt{2}}) \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} - (\tfrac{6}{\sqrt{8}}) \begin{bmatrix} \frac{2}{\sqrt{8}} \\ 0 \\ 0 \\ \frac{-2}{\sqrt{8}} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = w_3$$

Step 5: Find e_3

$$e_3 = \frac{w_3}{|w_3|} \to |w_3| = \sqrt{0^2 + 1^2 + 1^2 + 0^2} = \sqrt{2}$$

$$e_3 = \frac{w_3}{\sqrt{2}} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = e_3$$

An orthonormal basis for V is:

$$\{e_1, e_2, e_3\} = \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{2}{\sqrt{8}} \\ 0 \\ 0 \\ \frac{-2}{\sqrt{8}} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \right\}$$

That is the Gram Schmidt process for finding orthonormal bases. This can be done for any vector spaces and subspaces, like this example being a subspace of $\mathbb{R}^4$.

These types of problems lend themselves to small errors in arithmetic leading to very incorrect final answers, so being very careful and also physically writing down each step helps with partial credit and retracing steps.

Projections and Gram Schmidt are used heavily in math, primarily with computer computation. This course also details how you can do **least squares regression** by hand using projections and orthonormal vectors, however this isn't traditionally tested, but a really cool application of these techniques.