Gram Schmidt Process

Before we begin, let’s remind ourselves that the following subspaces are equivalent for any pairs of linearly independent vectors \( \vec{v}_1, \vec{v}_2 \):

- \( \text{span}(\vec{v}_1, \vec{v}_2) \)
- \( \text{span}(\vec{v}_1, \alpha \vec{v}_2) \)
- \( \text{span}(\vec{v}_1, \vec{v}_1 + \vec{v}_2) \)
- \( \text{span}(\vec{v}_1, \vec{v}_1 - \vec{v}_2) \)
- \( \text{span}(\vec{v}_1, \vec{v}_2 - \alpha \vec{v}_1) \)

Now what should \( \alpha \) be if we would like \( \vec{v}_1 \) and \( \vec{v}_2 - \alpha \vec{v}_1 \) to be orthogonal to each other? Intuitively, \( \alpha \vec{v}_1 \) should be the projection of \( \vec{v}_2 \) onto \( \vec{v}_1 \). Let’s solve this algebraically using the definition of orthogonality:

\[
\vec{v}_1 \text{ and } \vec{v}_2 - \alpha \vec{v}_1 \text{ are orthogonal} \quad (1)
\]
\[
\Leftrightarrow \vec{v}_1^T (\vec{v}_2 - \alpha \vec{v}_1) = 0 \quad (2)
\]
\[
\Leftrightarrow \vec{v}_1^T \vec{v}_2 - \alpha \| \vec{v}_1 \|^2 = 0 \quad (3)
\]
\[
\Leftrightarrow \alpha = \frac{\vec{v}_1^T \vec{v}_2}{\| \vec{v}_1 \|^2} \quad (4)
\]

**Definition 17.1 (Orthonormal):** A set of vectors \( \{\vec{v}_1, \ldots, \vec{v}_n\} \) is orthonormal if all the vectors are mutually orthogonal to each other and all are of unit length.

Gram Schmidt is an algorithm that takes a set of linearly independent vectors \( \{\vec{v}_1, \ldots, \vec{v}_n\} \) and generates an orthonormal set of vectors \( \{w_1, \ldots, w_n\} \) that span the same vector space as the original set. Concretely, \( \{w_1, \ldots, w_n\} \) needs to satisfy the following:

- \( \text{span}(\{v_1, \ldots, v_n\}) = \text{span}(\{w_1, \ldots, w_n\}) \)
- \( \{w_1, \ldots, w_n\} \) is an orthonormal set of vectors

Now let’s see how we can do this with a set of three vectors \( \vec{v}_1, \vec{v}_2, \vec{v}_3 \) that is linearly independent of each other.
• **Step 1:** Find unit vector $\vec{w}_1$ such that $\text{span}\{\{\vec{w}_1\}\} = \text{span}\{\{\vec{v}_1\}\}$.

Since $\text{span}\{\{\vec{v}_1\}\}$ is a one dimensional vector space, the unit vector that spans the same vector space would just be the normalized vector point in the same direction as $\vec{v}_1$. We have

$$\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}. \quad (5)$$

• **Step 2:** Given $\vec{w}_1$ from the previous step, find $\vec{w}_2$ such that $\text{span}\{\{\vec{w}_1, \vec{w}_2\}\} = \text{span}\{\{\vec{v}_1, \vec{v}_2\}\}$ and orthogonal to $\vec{w}_1$. We know that $\vec{v}_2 - (\text{projection of } \vec{v}_2 \text{ on } \vec{w}_1)$ would be orthogonal to $\vec{w}_1$ as seen earlier. Hence, a vector $\vec{e}_2$ orthogonal to $\vec{w}_1$ where $\text{span}\{\{\vec{w}_1, \vec{e}_2\}\} = \text{span}\{\{\vec{v}_1, \vec{v}_2\}\}$ is

$$\vec{e}_2 = \vec{v}_2 - (\vec{v}_2^T \vec{w}_1) \vec{w}_1. \quad (6)$$

Normalizing, we have $\vec{w}_2 = \frac{\vec{e}_2}{\|\vec{e}_2\|}$.

• **Step 3:** Now given $\vec{w}_1$ and $\vec{w}_2$ in the previous steps, we would like to find $\vec{w}_3$ such that $\text{span}\{\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}\} = \text{span}\{\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}\}$. We know that the projection of $\vec{v}_3$ onto the subspace spanned by $\vec{w}_1, \vec{w}_2$ is

$$(\vec{v}_3^T \vec{w}_2) \vec{w}_2 + (\vec{v}_3^T \vec{w}_1) \vec{w}_1. \quad (7)$$

We know that

$$\vec{e}_3 = \vec{v}_3 - (\vec{v}_3^T \vec{w}_2) \vec{w}_2 - (\vec{v}_3^T \vec{w}_1) \vec{w}_1 \quad (8)$$

is orthogonal to $\vec{w}_1$ and $\vec{w}_2$. Normalizing, we have $\vec{w}_3 = \frac{\vec{e}_3}{\|\vec{e}_3\|}$.

We can generalize the above procedure for any number of linearly independent vectors as follows:

1. **Inputs:**
   - A set of linearly independent vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$.

2. **Outputs:**
   - An orthonormal set of vectors $\{\vec{w}_1, \ldots, \vec{w}_n\}$ where $\text{span}\{\{\vec{v}_1, \ldots, \vec{v}_n\}\} = \text{span}\{\{\vec{w}_1, \ldots, \vec{w}_n\}\}$.

3. **procedure** **GRAM SCHMIDT**($\vec{v}_1, \ldots, \vec{v}_n$)

4. $\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}$

5. for $i = 2 \ldots n$ do

6. $\vec{e}_i \leftarrow \vec{v}_i - \sum_{j=1}^{i-1} (\vec{v}_i^T \vec{e}_j) \vec{w}_j$

7. $\vec{w}_i \leftarrow \frac{\vec{e}_i}{\|\vec{e}_i\|}$

8. end for

9. **end procedure**