

1 Complex Inner Product

For the complex vector space \mathbb{C}^n , we can no longer use our conventional real dot product as a valid inner product for \mathbb{C}^n . This is because the real dot product is no longer positive-definite for complex vectors.

For example, let $\vec{v} = \begin{bmatrix} j \\ j \end{bmatrix}$. Then, $\vec{v} \cdot \vec{v} = j^2 + j^2 = -2 < 0$.

Therefore, for two vectors $\vec{u}, \vec{v} \in \mathbb{C}^n$, we define the complex inner product to be:

$$\langle \vec{u}, \vec{v} \rangle = \sum_{i=1}^n u_i \bar{v}_i = \vec{u}^T \bar{\vec{v}} = \bar{\vec{v}}^* \vec{u}$$

where $\bar{\cdot}$ denotes the complex conjugate and $\bar{\cdot}^*$ denotes the complex conjugate transpose. Recall that the complex conjugate of a complex number $z = a + jb = re^{j\theta}$ is equal to

$\bar{z} = a - jb = re^{-j\theta}$. The conjugate transpose of a vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ is $\vec{v}^* = [\bar{v}_1 \quad \bar{v}_2 \quad \cdots \quad \bar{v}_n]$.

Note that this inner product is no longer symmetric but conjugate-symmetric, i.e., $\langle \vec{u}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{u} \rangle}$.

2 Gram-Schmidt Process

Gram-Schmidt is an algorithm that takes a set of linearly independent vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ and generates an orthonormal set of vectors $\{\vec{u}_1, \dots, \vec{u}_n\}$ that span the same vector space as the original set. Concretely, $\{\vec{u}_1, \dots, \vec{u}_n\}$ satisfy the following:

- $\forall 0 < k \leq n$, $\text{span}(\{\vec{v}_1, \dots, \vec{v}_k\}) = \text{span}(\{\vec{u}_1, \dots, \vec{u}_k\})$
- $\{\vec{u}_1, \dots, \vec{u}_n\}$ is an orthonormal set of vectors

Orthonormal Vectors

A set of vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ is **orthonormal** if all the vectors are mutually orthogonal to each other and all are of unit length. That is:

- Orthogonal: For all pairs of vectors \vec{v}_i, \vec{v}_j where $i \neq j$, $\langle \vec{v}_i, \vec{v}_j \rangle = 0$. For real vectors, this means $\vec{v}_i^T \vec{v}_j = 0$.
- Normalized: For all i , $\|\vec{v}_i\| = 1$. (This implies that $\|\vec{v}_i\| = \langle \vec{v}_i, \vec{v}_i \rangle = 1$.)

Projection of a Vector

The projection of a vector \vec{w} onto another vector \vec{v} is denoted as $\text{proj}_{\vec{v}}(\vec{w})$

$$\text{proj}_{\vec{v}}(\vec{w}) = \frac{\langle \vec{w}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v} = \frac{\langle \vec{w}, \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v} \quad (1)$$

If we project onto a vector \vec{v} that is normalized, $\|\vec{v}\| = 1$, then the projection will be

$$\text{proj}_{\vec{v}}(\vec{w}) = \langle \vec{w}, \vec{v} \rangle \vec{v} \quad (2)$$

Gram-Schmidt Algorithm

The Gram-Schmidt algorithm works by first finding the unit vector \vec{u}_1 such that $\text{span}(\{\vec{u}_1\}) = \text{span}(\{\vec{v}_1\})$. Subsequently, the unit vector \vec{u}_2 is calculated such that $\langle \vec{u}_1, \vec{u}_2 \rangle = 0$ and $\text{span}(\{\vec{u}_1, \vec{u}_2\}) = \text{span}(\{\vec{v}_1, \vec{v}_2\})$. This is continued through n vectors, resulting in the orthonormal set of vectors $\{\vec{u}_1, \dots, \vec{u}_n\}$ that span the same vector space as $\{\vec{v}_1, \dots, \vec{v}_n\}$.

How is this done? Finding \vec{u}_1 is straightforward, since it is the first vector in our new set, and therefore we must only satisfy $\|\vec{u}_1\| = 1$ and $\text{span}(\{\vec{u}_1\}) = \text{span}(\{\vec{v}_1\})$. 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 unit vector in the same direction as \vec{v}_1 . We have

$$\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}. \quad (3)$$

Calculating \vec{u}_2 requires that we satisfy:

1. Spanning the same vector space as original set: $\text{span}(\{\vec{u}_1, \vec{u}_2\}) = \text{span}(\{\vec{v}_1, \vec{v}_2\})$
2. Orthogonal to previous vectors: $\langle \vec{u}_1, \vec{u}_2 \rangle = 0$
3. Normalized: $\|\vec{u}_2\| = 1$

Using the vector \vec{u}_1 that we calculated above, we notice that

$$\text{span}(\{\vec{u}_1, \vec{v}_2\}) = \text{span}(\{\vec{v}_1, \vec{v}_2\}),$$

satisfying the first condition. However, \vec{u}_1 and \vec{v}_2 are not necessarily orthogonal. Recall from 16A/54 that the following subspaces are equivalent for any pair of linearly independent vectors \vec{v}_1, \vec{v}_2 :

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

Therefore, let us pick

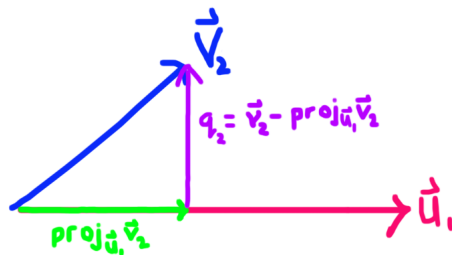
$$\vec{q}_2 = \vec{v}_2 - \alpha \vec{u}_1,$$

so that $\text{span}(\vec{u}_1, \vec{q}_2)$ will have the same span as $\{\vec{u}_1, \vec{v}_2\} = \text{span}(\vec{v}_1, \vec{v}_2)$.

What should α be if we would like \vec{u}_1 and $\vec{q}_2 = \vec{v}_2 - \alpha \vec{u}_1$ to be orthogonal to each other? We can recall from 16A, that subtracting the projection $\text{proj}_{\vec{u}_1}(\vec{v}_2)$ will create a vector orthogonal to \vec{u}_1 , where

$$\text{proj}_{\vec{u}_1}(\vec{v}_2) = \langle \vec{v}_2, \vec{u}_1 \rangle \vec{u}_1$$

is the projection of \vec{v}_2 onto \vec{u}_1 .



This makes sense because the projection of \vec{v}_2 onto \vec{u}_1 provides the component of \vec{v}_2 that is along \vec{u}_1 . Subtracting off this component from \vec{v}_2 will only leave components of \vec{v}_2 that are orthogonal to \vec{u}_1 .

Therefore, if we set

$$\alpha \vec{u}_1 = \text{proj}_{\vec{u}_1}(\vec{v}_2),$$

the resulting

$$\vec{q}_2 = \vec{v}_2 - \text{proj}_{\vec{u}_1}(\vec{v}_2) = \vec{v}_2 - \langle \vec{v}_2, \vec{u}_1 \rangle \vec{u}_1$$

will be orthogonal to \vec{u}_1 .

Finally, we normalize \vec{q}_2 to complete the process of finding the \vec{u}_2 which satisfies all three conditions above:

$$\vec{u}_2 = \frac{\vec{q}_2}{\|\vec{q}_2\|}$$

In the question below, you will work through how this methodology leads to the Gram-Schmidt algorithm for calculating the orthonormal set $\{\vec{u}_1, \dots, \vec{u}_n\}$ from n linearly independent vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$. Note that the Gram-Schmidt process can be done for the real or complex inner product so we will leave all of our results in general inner product form.

3 Gram-Schmidt Algorithm

Let's apply Gram-Schmidt orthonormalization to a set of three linearly independent vectors $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

- a) Find unit vector \vec{u}_1 such that $\text{span}(\{\vec{u}_1\}) = \text{span}(\{\vec{v}_1\})$.
- b) Given \vec{u}_1 from the previous step, find \vec{u}_2 such that $\text{span}(\{\vec{u}_1, \vec{u}_2\}) = \text{span}(\{\vec{v}_1, \vec{v}_2\})$ and \vec{u}_2 is orthogonal to \vec{u}_1 .

c) Now given \vec{u}_1 and \vec{u}_2 in the previous steps, find \vec{u}_3 such that $\text{span}(\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}) = \text{span}(\{\vec{v}_1, \vec{v}_2, \vec{v}_3\})$.

d) Let's extend this algorithm to n linearly independent vectors. That is, given an input $\{\vec{v}_1, \dots, \vec{v}_n\}$, write the algorithm to calculate the orthonormal set of vectors $\{\vec{u}_1, \dots, \vec{u}_n\}$, where $\text{span}(\{\vec{v}_1, \dots, \vec{v}_n\}) = \text{span}(\{\vec{u}_1, \dots, \vec{u}_n\})$. *Hint: How would you calculate the i^{th} vector, \vec{q}_i ?*

4 The Order of Gram-Schmidt

- a) If we are performing the Gram-Schmidt method on a set of vectors, does the order in which we take the vectors matter? Consider the set of vectors

$$\left\{ \vec{v}_1, \vec{v}_2, \vec{v}_3 \right\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad (4)$$

Perform Gram-Schmidt on these vectors first in the order $\vec{v}_1, \vec{v}_2, \vec{v}_3$.

- b) Now perform Gram-Schmidt on these vectors in the order $\vec{v}_3, \vec{v}_2, \vec{v}_1$. Do you get the same result?