Linear Algebra
(Inner products and orthonormal bases)

Mohamed El-Refai
ID: 900222334

Abdallah W. Mahmoud
ID: 900221058

The American University in Cairo

Fall 2023

"Mathematics is the language in which God has written the universe"
Galileo Galilei

Abstract.

In this paper, we introduced an important topic that appears in both linear algebra and calculus, and provides mathematicians and physicists alike with a new perspective on vectors. We first begin with inner product spaces, where we discussed it's application through projections, and then a detailed exploration of orthonormal bases. Also, we reinforced these concepts by going into the Gram-Schmidt process. Our goal is to provide an complete understanding of inner product spaces and their associated operations. Enjoy this journey of learning!

1   Motivation

Once upon a time in Math land, where numbers had their own neighborhood and variables were like friendly neighbors, two special numbers, Omar and Ali, wanted to measure how close they were. They weren't sure how to do it until they discovered a magical tool called "inner product spaces" What are inner product spaces? Well, imagine you have a favorite song. Each instrument in the song plays a unique part, right? Inner product spaces help us understand how these different musical elements come together perfectly.

But, hold on, this isn't just about music. Inner product spaces help us in all kinds of areas, like figuring out directions, understanding how signals work in your phone, and even diving into the weird world of quantum stuff.

In this paper, we'll explore Abstract Inner Products and Orthogonal Projections. Abstract Inner Products provide a unique way to understand geometric relationships in a broader mathematical context, acting as a language for describing connections beyond traditional rules. Also, we'll dive into Orthogonal Projections, a tool that helps break down vectors into different parts, showing us their real-world applications. By the end of this journey, readers will know these fundamental concepts, both in theory and in practical scenarios, offering an understanding of Abstract Inner Products and the utility of Orthogonal Projections.

2   The Familiar Case

Let's begin by considering the familiar concept of the dot product in $\mathbb{R}^n$ as an example. In $\mathbb{R}^n$, the dot product serves as a fundamental operation.

Firstly, let's first explore the case when $n = 2$ or $n = 3$. In $\mathbb{R}^2$, suppose we have two vectors, $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$. The dot product is computed as:

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2$$

Similarly, in $\mathbb{R}^3$, with vectors $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$, the dot product is given by:

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + v_3 w_3$$

Now, why is this dot product so essential? It provides us with a means to compute crucial geometric properties in $\mathbb{R}^n$, such as the length (or magnitude) of a vector, the angle between vectors, and the distance between points.

Vector Length: The length ($|\mathbf{v}|$) of a vector $\mathbf{v}$ in $\mathbb{R}^2$ is calculated using the dot product:

$$|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$$

Angle Between Vectors: The angle (θ) between two vectors $\mathbf{v}$ and $\mathbf{w}$ is determined by:

$$\cos(\theta) = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}| \cdot |\mathbf{w}|}$$

Distance Between Points: The distance between two points $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^2$ is given by:

$$d(\mathbf{v}, \mathbf{w}) = |\mathbf{v} - \mathbf{w}| = \sqrt{(\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w})}$$

Having explored these fundamental concepts in $\mathbb{R}^2$ and $\mathbb{R}^3$, we can now generalize these notions to any dimension n. The dot product not only facilitates these calculations but also lays the groundwork for the broader understanding of inner product spaces, where these operations can be extended and generalized in a more abstract and powerful mathematical framework.

3 Abstract Inner Products

We start by defining the inner product within an arbitrary vector space $(V, +, \cdot)$.

Definition 3.1. Let $(V, +, \cdot)$ be a vector space. An **inner product** on V is a function $\langle , \rangle$ on V that maps a pair of vectors $\mathbf{u}, \mathbf{v} \in V$ to a scalar $\langle \mathbf{u}, \mathbf{v} \rangle$. As we know, a function is essentially a linear transformation, so $V \times V \to \mathbb{R}$

via $\langle , \rangle$. This function should satisfy the following properties.

Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\alpha \in \mathbb{R}$. Then:

(IP1) **Commutativity**
$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \; \forall \; \mathbf{u}, \mathbf{v} \in V$
In fact, this property is part of a larger property called **conjugate symmetry** but because conjugate symmetry deals with complex vector spaces, we use commutativity instead for real vector spaces.

(IP2) **Positivity**
$\langle \mathbf{u}, \mathbf{u} \rangle \geq 0 \; \forall \; \mathbf{u} \in V$

(IP3) **Definiteness**
$\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \underline{\mathbf{0}}$

(IP4) **Compatibility**
$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ and $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle \; \forall \; \mathbf{u}, \mathbf{v} \in V$ and $\forall \; \alpha \in \mathbb{R}$

Example 3.2 (The standard inner product on $\mathbb{R}^n$). *Consider V to be the set $\mathbb{R}^n$, and consider two vectors $\underline{\mathbf{u}} = (x_1, x_2, \ldots, x_n)$ and $\underline{\mathbf{v}} = (y_1, y_2, \ldots, y_n)$. Then we define the standard inner product on V via $\langle u, v \rangle := x_1 y_1 + x_2 y_2 + \ldots, + x_n y_n$ $= \sum_{i=1}^{n} x_i y_i$*

(A non standard inner product) Consider V to be the set $\mathbb{R}^2$, and consider two vectors $\underline{\mathbf{u}} = (x_1, y_1)$ and $\underline{\mathbf{v}} = (x_2, y_2)$. We define the product on V via $\langle u, v \rangle := 2x_1 x_2 + 4y_1 y_2$. So if $\underline{\mathbf{u}} = (1, 2)$ and $\underline{\mathbf{v}} = (3, 4)$ then $\langle u, v \rangle := 2(1)(3) + 4(4)(2) = 38$ We now need to test if the function is an inner product or not. To do that, we test if the four properties are satisfied.

(IP1) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 2x_1 x_2 + 4y_1 y_2 = 2x_2 x_1 + 4y_2 y_1 = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$ so IP1 satisfied

(IP2) $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = \sqrt{x_1^2 + y_1^2} \geq 0$
Also, $\langle v, v \rangle = \sqrt{x_2^2 + y_2^2} \geq 0$ so IP2 satisfied

(IP3) if we assume that $\boldsymbol{u} = \underline{\boldsymbol{0}}$, then $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = \sqrt{0 + 0} = 0$ and we similarly do that for $\boldsymbol{v}$

(IP4) let $\boldsymbol{w} = (x_3, y_3)$ ($\boldsymbol{w} \in \mathbb{R}^2$) then $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} \rangle = \langle (x_1 + x_2, y_1 + y_2), (x_3, y_3) \rangle$

     $= 2x_3(x_1 + x_2) + 4y_3(y_1 + y_2)$

     $= 2x_3x_1 + 2x_3x_2 + 4y_3y_1 + 4y_3y_2 = 2x_3x_1 + 4y_3y_1 + 2x_3x_2 + 4y_3y_2$

     $= ((x_3, y_3), (x_1, y_1)) + ((x_3, y_3), (x_2, y_2)) = \langle \boldsymbol{w}, \boldsymbol{u} \rangle + \langle \boldsymbol{w}, \boldsymbol{v} \rangle$

     And we know from IP1 that $\langle \boldsymbol{w}, \boldsymbol{u} \rangle = \langle \boldsymbol{u}, \boldsymbol{w} \rangle$ and $\langle \boldsymbol{w}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$

     Also, consider $\alpha \in \mathbb{R}$, so $\langle \alpha \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \alpha(x_1, y_1), (x_2, y_2) \rangle = 2\alpha x_1 x_2 + 4\alpha y_1 y_2$

     $= \alpha(2x_1x_2 + 4y_1y_2) = \alpha \langle \boldsymbol{u}, \boldsymbol{v} \rangle$

     so IP4 also satisfied. Hence the function $\langle u, v \rangle := 2x_1x_2 + 4y_1y_2$ is an inner product.

Lemma 3.3 (A generalised dot product)**.** *For any $a, b \in \mathbb{R}$ with $a, b > 0$, the function $\langle \cdot, \cdot \rangle_{a,b} : \mathbb{R}^2 \to \mathbb{R}^2$ defined via*

$$\langle (x_1, y_1), (x_2, y_2) \rangle_{a,b} := ax_1x_2 + by_1y_2$$

defines an inner product on the vector space $(\mathbb{R}^2, +, \cdot)$.

Proof. let $\underline{\mathbf{u}} = (x_1, y_1)$ and $\underline{\mathbf{v}} = (x_2, y_2)$, and $\underline{\mathbf{u}}, \underline{\mathbf{v}} \in \mathbb{R}^2$, then we get
$\langle \underline{\mathbf{u}}, \underline{\mathbf{v}} \rangle_{ab} = a(x_1x_2) + b(y_1y_2) = ax_1x_2 + by_1y_2$ ∎

Note that the axioms in the definition of an IPS are independent, i.e. neither one follows from the others. To illustrate this, we will give examples of functions which satisfy some, but not all of the properties of an IPS.

Example 3.4. Let $\mathbf{u} = (u_1, u_2)$, $\mathbf{v} = (v_1, v_2)$, $\mathbf{w} = (w_1, w_2)$, $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^2$
The following are non-examples of inner products.

1. Consider the function $\langle \mathbf{u}, \mathbf{v} \rangle = u_1^2 v_1^2 + u_2^2 v_2^2$. This function satisfies IP1, IP2 and IP3 but does not satisfy IP4.

2. Consider the function $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 - u_2 v_2$. This function satisfies IP1 and IP4 but does not satisfy IP2 and IP3

3. Consider the function $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_2 + u_2 v_1$. This function satisfies IP1 and IP4 but does not satisfy IP2 and IP3

Now consider $\mathbf{u} = (u_1, u_2, \ldots, u_n), \mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$, then:

1. we define the norm of $\mathbf{u}$ to be $\sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = ||\mathbf{u}||$ and likewise for $\mathbf{v}$.

2. we define the angle (θ) between two vectors $\mathbf{u}$ and $\mathbf{v}$ is determined by:

$$\cos(\theta) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{u}|| \cdot ||\mathbf{v}||}$$

If the angle θ between $\mathbf{u}$ and $\mathbf{v}$ is 0, then we simplify the equation to $\langle \mathbf{u}, \mathbf{v} \rangle = ||\mathbf{u}|| \cdot ||\mathbf{v}||$
And if the angle θ between $\mathbf{u}$ and $\mathbf{v}$ is 180 degrees, then we simplify the equation to $\langle \mathbf{u}, \mathbf{v} \rangle = -||\mathbf{u}|| \cdot ||\mathbf{v}||$
And if the angle θ between $\mathbf{u}$ and $\mathbf{v}$ is 90 degrees, then they are orthogonal, then we simplify the equation to $||\mathbf{u}|| \cdot ||\mathbf{v}|| = \langle \mathbf{u}, \mathbf{v} \rangle = 0$

3. we define the distance between two points $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is given by:

$$d(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}|| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$$

While the generalized dot product yields inner products for all standard spaces $\mathbb{R}^n$, the general definition of an inner product now also allows us to talk about distances of functions, polynomials and matrices. The following examples exhibit the standard inner products on these vector spaces.

Example 3.5 (An inner product on $M_{m \times n}$). *Let $A, B \in \mathbb{M}_{m*n}$ and a_{ij} and b_{ij} be random elements in A and B respectively, where i can vary from 0 to m, and j can vary from 0 to n. Then*
$\langle A, B \rangle = a_{11}b_{11} + a_{12}b_{12} + \cdots + a_{mn}b_{mn}$

Example 3.6 (An inner product on $\mathcal{P}_{\leq n}$). *Let $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ and $q(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n$. Then*
$\langle p(x), q(x) \rangle = a_0 b_0 + a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$

Example 3.7 (An inner product on continuous functions). *Consider the functions $f(x)$ and $g(x)$ which is continuous on the interval [-1,1]. Then*
$\langle f(x), g(x) \rangle = \int_{-1}^{1} f(x) \cdot g(x) \, dx$

Lemma 3.8. *In any inner product space $(V, \langle \cdot \rangle)$, the vector $\underline{0}$ is orthogonal to any other vector $v \in V$.*

Proof. Consider $\alpha \in R$. We know from IP4 that $\langle \alpha\mathbf{u}, \mathbf{v}\rangle = \alpha\langle\mathbf{u}, \mathbf{v}\rangle$, and let $\mathbf{u}=\underline{\mathbf{0}}$. Then $\alpha * \mathbf{u} = \underline{\mathbf{0}} \; \forall \; \alpha \in R$ so $\langle\alpha\mathbf{u}, \mathbf{v}\rangle = \alpha\langle\mathbf{u}, \mathbf{v}\rangle = \underline{\mathbf{0}} \forall$ u,v$\in V$ Hence, any vector $\mathbf{v}\in V$ is orthogonal to the zero vector. ∎

Theorem 3.9 (Cauchy-Schwarz Inequality)**.** *Consider two vectors $\boldsymbol{u}, \boldsymbol{v}$ in V. Then $|\langle\boldsymbol{u}, \boldsymbol{v}\rangle| \leq ||u||||v||$*

Proof. Assume $\mathbf{v}, \mathbf{u} \neq \underline{\mathbf{0}}$ and let α *in* $\mathbb{R}$. Then, $\langle\mathbf{u}-\alpha\mathbf{v}, \mathbf{u}-\alpha\mathbf{v}\rangle = ||\langle\mathbf{u}-\alpha\mathbf{v}|| \geq 0$
$= ||u||^2 - 2\alpha\langle\mathbf{u}, \mathbf{v}\rangle + \alpha^2||v||^2 \geq 0$
Then let $\alpha = \frac{\langle\mathbf{u},\mathbf{v}\rangle}{||\mathbf{u}||^2 \cdot ||\mathbf{v}||^2}$. Substituting α into the inequality, we get $\frac{\langle\mathbf{u},\mathbf{v}\rangle^2}{||\mathbf{v}||^2} \leq ||\mathbf{u}||^2$. After removing the squares, we get $|\langle\mathbf{u}, \mathbf{v}\rangle| \leq ||u||||v||$, proving the theorem. ∎

Theorem 3.10 (Triangle Inequality)**.** *The Cauchy-Shwarz Inequality leads us to the triangle inequality, given by $||\boldsymbol{u} + \boldsymbol{v}|| \leq ||\boldsymbol{u}|| + ||\boldsymbol{v}||$*
Proof : We compute $||\langle\boldsymbol{u} + \boldsymbol{v}\rangle||^2$ this time instead of $||\langle\boldsymbol{u}, \boldsymbol{v}\rangle||^2$ and from the Cauchy-Shwarz inequality, we get
$||\boldsymbol{u} + \boldsymbol{v}||^2 \leq ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2 + 2||\boldsymbol{u}||||\boldsymbol{v}||$ this leads up to
$||\boldsymbol{u} + \boldsymbol{v}||^2 \leq ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2$ which simplifies to
$||\boldsymbol{u} + \boldsymbol{v}|| \leq ||\boldsymbol{u}|| + ||\boldsymbol{v}||$

Corollary 3.11 (Theorem of Pythagoras)**.** *Let $\langle\boldsymbol{u}, \boldsymbol{v}\rangle = \underline{\boldsymbol{0}}$ meaning that $\boldsymbol{u}$ is orthogonal to $\boldsymbol{v}$. Then*
$||\boldsymbol{u} + \boldsymbol{v}||^2 = ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2$
Proof: We know that $||\boldsymbol{u} + \boldsymbol{v}||^2 = \langle\boldsymbol{u} + \boldsymbol{v}, \boldsymbol{u} + \boldsymbol{v}\rangle$, and we know from IP4 that
$\langle\boldsymbol{u} + \boldsymbol{v}, \boldsymbol{u} + \boldsymbol{v}\rangle = \langle\boldsymbol{u}, \boldsymbol{u}\rangle + \langle\boldsymbol{u}, \boldsymbol{v}\rangle + \langle\boldsymbol{v}, \boldsymbol{u}\rangle + \langle\boldsymbol{v}, \boldsymbol{v}\rangle$
Using IP1, we simplify that to $||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2 + 2||\boldsymbol{u}||||\boldsymbol{v}||$ and we know that
$\langle\boldsymbol{u}, \boldsymbol{v}\rangle = \underline{\boldsymbol{0}}$.
Hence, we get $||\boldsymbol{u} + \boldsymbol{v}||^2 = ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2$ as desired.

Now that we are familiar with manipulating inner products and computing the angles between vectors and their norms, we move on to the idea of projecting a vector which belongs in a certain vector space.

4 Orthogonal Projections

Orthogonal projections goes beyond measuring distances solely between vectors and it introduces an important concept. Instead of just understanding

how far one point is from another point, we understand how far a point is from a whole set of points, known as a subspace. Orthogonal projections are crucial in fields like graphics and machine learning where understanding the closeness of data points to specific groups is essential for making decisions.

Example 4.1. *Let's consider the example in $\mathbb{R}^2$, where we have vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. The orthogonal projection of $\mathbf{u}$ onto $\mathbf{v}$ is given by the formula:*

$$proj_{\mathbf{v}}\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v}$$

In this formula, $proj_{\mathbf{v}}\mathbf{u}$ is the orthogonal projection of $\mathbf{u}$ onto $\mathbf{v}$, $\langle \mathbf{u}, \mathbf{v} \rangle$ is the dot product of $\mathbf{u}$ and $\mathbf{v}$, and $\langle \mathbf{v}, \mathbf{v} \rangle$ is the dot product of $\mathbf{v}$ with itself.

The resulting vector $proj_{\mathbf{v}}\mathbf{u}$ lies on the line spanned by $\mathbf{v}$ and is special because it is the closest vector to $\mathbf{u}$ on that line. It represents the component of $\mathbf{u}$ that aligns with the direction of $\mathbf{v}$.

4.1   Orthonormal Bases

The Orthonormal Bases introduces an important concept that simplifies math in linear algebra, in which they make complex operations easier and gives us a clearer picture of transformations and spaces.

Definition 4.2. *A set of vectors is considered orthogonal if all pairs of vectors in the set are perpendicular which also means having dot product of zero. If the set is orthogonal and the vectors have a length of 1, then the set is orthonormal.*

Example 4.3. *Two examples of orthonormal bases for $\mathbb{R}^3$ are:*

1. Standard Basis: The standard basis for $\mathbb{R}^3$ consists of the following three vectors:

$$\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

This basis is orthonormal because each vector is orthogonal to the others, and they all have a length (norm) of 1.

2. Unit Vectors along Axes: Another orthonormal basis for $\mathbb{R}^3$ can be constructed using unit vectors along the coordinate axes:

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

These vectors are also mutually orthogonal and have a length of 1, forming an orthonormal basis.

Example 4.4. *ONB for $M_{m \times n}$: Let's consider the space of 2×2 matrices. An orthonormal basis can be formed using the following matrices:*

$$\mathbf{M}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{M}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{M}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{M}_4 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

These matrices form an orthonormal basis for $M_{2 \times 2}$ because they are mutually orthogonal and each has a Frobenius norm of 1.

ONB for $P_{\leq n}$: For the space of polynomials of degree less than or equal to n, an orthonormal basis can be constructed using the Legendre polynomials. For $n = 2$, the Legendre polynomials are $P_0(x) = 1$, $P_1(x) = x$, and $P_2(x) = \frac{1}{2}(3x^2 - 1)$. To normalize them, we divide each by its norm:

$$p_0(x) = \frac{1}{\sqrt{2}}, \quad p_1(x) = \sqrt{\frac{3}{2}}x, \quad p_2(x) = \frac{1}{2}\sqrt{\frac{5}{2}}(3x^2 - 1)$$

These polynomials are mutually orthogonal and form an orthonormal basis for $P_{\leq 2}$.

Example 4.5. *To show that the set $S := \{1, \sin(x), \cos(x), \sin(2x), \cos(2x), \ldots\}$ is orthogonal in the space $(C([-1, 1]), \langle \cdot, \cdot \rangle)$, we consider functions $f_n(x) = \sin(nx)$ and $g_m(x) = \cos(mx)$ for $n, m \in \mathbb{N}$. The inner product $\langle f_n, g_m \rangle$ is given by:*

$$\langle f_n, g_m \rangle = \int_{-1}^{1} f_n(x) g_m(x) \, dx$$

For $n \neq m$, f_n and g_m are orthogonal since the sin and cos have orthogonal properties. The integral of their product over the interval $[-1, 1]$ is zero, showing that S is an orthogonal set.

However, S is not orthonormal because the functions in S do not have unit norm. For example, the norm of $\sin(x)$ over the interval $[-1, 1]$ is not 1. To see this, calculate:

$$\langle \sin(x), \sin(x) \rangle = \int_{-1}^{1} \sin^2(x)\, dx$$

This integral is not equal to 1, indicating that $\sin(x)$ does not have unit norm.

In summary, S is orthogonal but not orthonormal in the space $(C([-1, 1]), \langle \cdot, \cdot \rangle)$.

Theorem 4.6. *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and $S := \{v_1, \ldots, v_n\} \subseteq V$ be an orthogonal set of nonzero vectors in V. Then S is linearly independent.*

Proof. Suppose there exist scalars $c_1, \ldots, c_n$ (not all zero) such that:

$$c_1 v_1 + c_2 v_2 + \ldots + c_n v_n = 0$$

Now, take the inner product of both sides with each vector v_i in S:

$$\langle c_1 v_1 + c_2 v_2 + \ldots + c_n v_n, v_i \rangle = \langle 0, v_i \rangle$$

Using the linearity of the inner product, this simplifies to:

$$c_1 \langle v_1, v_i \rangle + c_2 \langle v_2, v_i \rangle + \ldots + c_i \langle v_i, v_i \rangle + \ldots + c_n \langle v_n, v_i \rangle = 0$$

However, since S is an orthogonal set, $\langle v_j, v_i \rangle = 0$ for $i \neq j$. Therefore, the equation becomes:

$$c_i \langle v_i, v_i \rangle = 0$$

Since v_i is nonzero, $\langle v_i, v_i \rangle$ is also nonzero. Therefore, we must have $c_i = 0$. This holds for all i, and consequently, all coefficients in the linear combination are zero.

Thus, S is linearly independent. This completes the proof of the theorem.

$\blacksquare$

Example 4.7. *Let $S = \{(2, 2, -2), (1, 0, 1), (-1, 2, 1)\}$ be a set of vectors in $\mathbb{R}^3$. We want to show that S forms a basis for $\mathbb{R}^3$.*

Denote the vectors in S as v_1, v_2, v_3:

$$v_1 = (2, 2, -2), \quad v_2 = (1, 0, 1), \quad v_3 = (-1, 2, 1)$$

Now, let's check if S is an orthogonal set:

$$\langle v_1, v_2 \rangle = 2(1) + 2(0) + (-2)(1) = 0$$
$$\langle v_1, v_3 \rangle = 2(-1) + 2(2) + (-2)(1) = 0$$
$$\langle v_2, v_3 \rangle = (1)(-1) + (0)(2) + (1)(1) = 0$$

Since the inner product of any two distinct vectors in S is zero, S is an orthogonal set.

By the previous theorem, since S is an orthogonal set of nonzero vectors, it is also linearly independent. Since $\dim(\mathbb{R}^3) = 3$, and S has 3 vectors, S is a basis for $\mathbb{R}^3$.

Definition 4.8. *The coordinates of a vector with respect to a given basis represent the scalars needed to express that vector as a linear combination of the basis vectors. For a basis $\{v_1, v_2, \ldots, v_n\}$ and a vector v, the coordinates are the scalars $\{c_1, c_2, \ldots, c_n\}$ such that $v = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$.*

Now, let's use orthogonal projection and inner products to compute the coordinates of a vector with respect to a given orthonormal basis. Suppose we have a vector v and an orthonormal basis $\{u_1, u_2, \ldots, u_n\}$. The coordinates are obtained by taking the inner product of v with each basis vector:

$$c_i = \langle v, u_i \rangle$$

Example: Consider $v = (3, -1, 2)$ and the orthonormal basis $B = \{u_1, u_2, u_3\}$ where:

$$u_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right), \quad u_2 = \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right), \quad u_3 = (0, 0, 1)$$

The coordinates of v with respect to B are:

$$c_1 = \langle v, u_1 \rangle = \frac{3}{\sqrt{2}}, \quad c_2 = \langle v, u_2 \rangle = \frac{4}{\sqrt{3}}, \quad c_3 = \langle v, u_3 \rangle = 2$$

So, the coordinates of v with respect to B are $\left(\frac{3}{\sqrt{2}}, \frac{4}{\sqrt{3}}, 2 \right)$.

Now, let's use coordinate representation to prove that $w = (1, -2, 0, 1)$ is not in the span of $S = \left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right), \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0 \right), (0, 0, 0, 1) \right\}$. Suppose w is in the span of S, then the coordinates of w with respect to the basis S would exist. However, if we attempt to find these coordinates, we would find that w cannot be expressed as a linear combination of the vectors in S, indicating that w is not in the span of S.

4.2   The Gram-Schmidt Process

Imagine you have a set of vectors that are not orthogonal or normalized, and you wish to create a new set that is both orthogonal and normalized. This is precisely the reason behind the Gram-Schmidt Process. In linear algebra, transforming a set of vectors into an orthonormal basis is a fundamental operation, and the Gram-Schmidt Process provides a systematic way to achieve this.

The Gram-Schmidt Process is a method for transforming a linearly independent set of vectors into an orthogonal or orthonormal set. Given a set of vectors $\{v_1, v_2, \ldots, v_n\}$, the process proceeds as follows:

To start the process, we normalize the first vector, that is, we define

$$u_1 = \frac{s_1}{\|s_1\|}. \tag{1}$$

In the second step, we project s_2 on u_1:

$$s_2 = \langle s_2, u_1 \rangle u_1 + \varepsilon_2$$

where ε_2 is the residual of the projection. Then, we normalize the residual:

$$u_2 = \frac{\varepsilon_2}{\|\varepsilon_2\|}$$

The two vectors u_1 and u_2 thus obtained are orthonormal. In the third step, we project s_3 on u_1 and u_2:

$$s_3 = \langle s_3, u_1 \rangle u_1 + \langle s_3, u_2 \rangle u_2 + \varepsilon_3$$

and we compute the residual of the projection ε_3. We then normalize it:

$$u_3 = \frac{\varepsilon_3}{\|\varepsilon_3\|}$$

We proceed in this manner until we obtain the last normalized residual u_K. At the end of the process, the vectors $u_1, \ldots, u_K$ form an orthonormal set because:

- They are the result of a normalization, and as a consequence, they have unit norm.

- Each u_k is obtained from a residual that has the property of being orthogonal to $u_1, \ldots, u_{k-1}$.

It is important to recall that the linear span of $s_1, \ldots, s_K$ represents the collection of all vectors that can be expressed as combinations of $s_1, \ldots, s_K$. This span, denoted by $\text{span}(s_1, \ldots, s_K)$, constitutes a linear space. Given that the vectors $u_1, \ldots, u_K$ are independent linear combinations of $s_1, \ldots, s_K$, any vector that can be formed by combining $u_1, \ldots, u_K$ can equivalently be expressed as a combination of $s_1, \ldots, s_K$. Consequently, the spans of the two sets of vectors are identical: $\text{span}(s_1, \ldots, s_K) = \text{span}(u_1, \ldots, u_K)$.

Example 4.9. *Consider the set of vectors $\{(1, 0, 1), (2, 1, 1), (0, 1, -1)\}$. Apply the Gram-Schmidt Process to obtain an orthogonal set.*

According to the Gram-Schmidt process,

$$\mathbf{u}_k = \mathbf{v}_k - \sum_{j=1}^{k-1} \text{proj}(\mathbf{u}_j, \mathbf{v}_k)\mathbf{u}_j,$$

where

$$\text{proj}(\mathbf{u}_j, \mathbf{v}_k) = \frac{\mathbf{u}_j \cdot \mathbf{v}_k}{\|\mathbf{u}_j\|^2}\mathbf{u}_j$$

is a vector projection.

The normalized vector is

$$\mathbf{e}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.$$

Step 1

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{e}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \begin{bmatrix} \frac{2}{\sqrt{2}} \\ 0 \\ \frac{2}{\sqrt{2}} \end{bmatrix}$$

Step 2

$$\mathbf{u}_2 = \mathbf{v}_2 - \text{proj}(\mathbf{u}_1, \mathbf{v}_2)$$

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

$$\mathbf{e}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \begin{bmatrix} \frac{6}{\sqrt{6}} \\ \frac{6}{\sqrt{6}} \\ \frac{6}{\sqrt{6}} \end{bmatrix}$$

13

Step 3

$$\mathbf{u}_3 = \mathbf{v}_3 - \mathrm{proj}(\mathbf{u}_1, \mathbf{v}_3) - \mathrm{proj}(\mathbf{u}_2, \mathbf{v}_3)$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Since $\mathbf{u}_3$ is a zero vector, we skip it.

ANSWER The set of orthonormal vectors is

$$\left\{ \begin{bmatrix} \frac{2}{\sqrt{2}} \\ 0 \\ \frac{2}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{6}{\sqrt{6}} \\ \frac{6}{\sqrt{6}} \\ \frac{6}{\sqrt{6}} \end{bmatrix} \right\}$$

The Gram-Schmidt Process is a versatile tool applicable not only in Euclidean spaces like $\mathbb{R}^n$ but also in more abstract vector spaces, providing a powerful method for creating orthonormal bases.

5 Summary

In this paper, we discussed how to identify, compute and manipulate inner products. We also expanded on inner products by introducing projection and diving deep into orthonormal bases. We then reinforced these ideas by discussing the Gram-Schmidt process. Attached next page are the references we used for this paper. Happy learning!

References

[1] Sheldon Axler. *Linear Algebra Done Right, 2nd Edition*. Springer-Verlag, 2004.

[2] Eslam Badr. *Lecture notes.*

[3] eMathHelp. *Gram-Schmidt Calculator.*

[4] R. Larson. *Elementary Linear Algebra*. Cengage Learning, 2016.

[5] Isabel Muller. *Lecture notes.*

[6] Gilbert Strang. *Introduction to Linear Algebra, 5th edition*. Wellesly - Cambridge Press, 2016.

[7] Marco Toboga. *"Gram-Schmidt process", Lectures on matrix algebra.* 2021.

[1] [4] [6] [7] [3] [2] [5]