

## MFML SEC9

This guide is designed to build a deep, intuitive understanding of the mathematical tools used to compare data. We will move slowly and deliberately, focusing on the "why" behind the mathematics, using detailed examples and analogies to make the concepts clear.

# Contents

---

<b>1</b>	<b>The Quest for "Similarity"</b>	<b>2</b>
<b>2</b>	<b>The Standard Toolkit: A Deep Dive into the Dot Product</b>	<b>3</b>
2.1	Property 1: Linearity . . . . .	3
2.2	Property 2: Symmetry . . . . .	3
2.3	Property 3: Positive-Definiteness . . . . .	4
<b>3</b>	<b>Forging a New Toolkit: The General Inner Product</b>	<b>5</b>
<b>4</b>	<b>The Engine Room: Symmetric Positive-Definite Matrices</b>	<b>6</b>
4.1	The Dot Product's "Engine": The Identity Matrix . . . . .	6
4.2	Creating Custom Inner Products with New Matrices . . . . .	6
4.3	Why Positive-Definiteness is Essential: A Counterexample . . . . .	8
4.4	The Geometric Interpretation: How Matrix A Reshapes Space . . . . .	8
<b>5</b>	<b>What Our New Tools Can Do</b>	<b>10</b>
5.1	Custom Lengths and Distances . . . . .	10
5.2	Redefining Angles and Orthogonality . . . . .	10
<b>6</b>	<b>The Art of Simplification: Orthonormal Bases</b>	<b>12</b>
<b>7</b>	<b>Conclusion: A New Perspective on Geometry</b>	<b>14</b>

# 1 The Quest for "Similarity"

---

Imagine a service like Netflix or Spotify. How does it recommend a new movie or song you might like? At its core, it's trying to solve a problem of similarity. It represents you and all the items you've liked as data points—vectors—in a high-dimensional space. To make a recommendation, it searches for other items (songs, movies) whose vectors are "close" to yours.

But what does "close" actually mean?

Our default understanding of closeness is the straight-line distance we learned in school, often called the **Euclidean distance**. This is a perfectly valid way to measure distance, but it's not the only way.

What if, for comparing movies, the "director" was a much more important feature than the "runtime"? We might want a custom distance measure that gives more weight to the director. The standard ruler isn't enough; we need to design a new one.

This guide is about how to build those new rulers. Our journey will unfold in four parts:

1. **Standard Ruler:** We'll take a deep look at the dot product, the engine behind Euclidean distance, and uncover its fundamental properties.
2. **New Toolkit:** We'll use those properties as a blueprint to define a more powerful and flexible tool: the general inner product.
3. **Usability:** We'll see how any inner product gives us a new way to define length, distance, and even angles.
4. **Simplification:** We'll learn how to create ideal, clean coordinate systems for any of these new geometries using the Gram-Schmidt process.

## 2 Dot Product

---

Our investigation begins with the familiar dot product. To build something new, we must first understand the foundations of what we already have.

### Definition 1. Dot Product

For two vectors  $\mathbf{x} = [x_1, \dots, x_n]^T$  and  $\mathbf{y} = [y_1, \dots, y_n]^T$  in  $\mathbb{R}^n$ , their **dot product** is:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

More than just a formula, the dot product tells us about the alignment of two vectors. A positive result means they point in generally the same direction, a negative result means they point in generally opposite directions, and a zero result means they are perpendicular.

The true power of the dot product, however, lies in three fundamental properties that we often take for granted. These are the "rules of the game" that make it so useful.

### 2.1 Property 1: Linearity

The dot product distributes over vector addition and is compatible with scalar multiplication.

$$(a\mathbf{x} + b\mathbf{y}) \cdot \mathbf{z} = a(\mathbf{x} \cdot \mathbf{z}) + b(\mathbf{y} \cdot \mathbf{z})$$

#### Q Intuition

This means you can either combine your vectors first and then take the dot product, or take the dot products first and then combine the results. The outcome is the same. It's a well-behaved, distributive property.

### 2.2 Property 2: Symmetry

The order in which you take the dot product doesn't matter.

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$$

#### Q Intuition

This makes sense: the measure of similarity or relationship from vector  $\mathbf{x}$  to vector  $\mathbf{y}$  should be the same as the relationship from  $\mathbf{y}$  to  $\mathbf{x}$ .

## 2.3 Property 3: Positive-Definiteness

The dot product of a vector with itself is always positive, unless the vector is the zero vector.

$$\mathbf{x} \cdot \mathbf{x} \geq 0, \quad \text{and} \quad \mathbf{x} \cdot \mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$$

### ?

#### Intuition

This is perhaps the most important property for geometry. Since  $\mathbf{x} \cdot \mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2$ , this is just the squared length of the vector from the Pythagorean Theorem. This property states that:

- The length of any real object cannot be negative.
- The only object with zero length is an object of no size—the zero vector.

Without this property, we could not have a coherent concept of length or distance.

## 3 The General Inner Product

---

Here is where we make our creative leap. We've identified the three essential rules that the dot product follows.

**The Big Idea:** What if we use these three rules—Linearity, Symmetry, and Positive-Definiteness—as the *blueprint* for a whole class of new tools?

Any operation that satisfies this blueprint is called an **inner product**.

### Definition 2. Inner Product

An **inner product** on a real vector space  $V$  is a function  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$  that satisfies these three axioms:

1. **Bilinearity:** It is linear in both of its arguments.
2. **Symmetry:**  $\langle x, y \rangle = \langle y, x \rangle$  for all  $x, y \in V$ .
3. **Positive-Definiteness:**  $\langle x, x \rangle > 0$  for all  $x \neq 0$ .

A vector space paired with an inner product,  $(V, \langle \cdot, \cdot \rangle)$ , is called an inner product space.

This naturally leads to the most important practical question: if the dot product isn't the only inner product, how do we find or create others?

# 4 Symmetric Positive-Definite Matrices

The secret recipe for creating new inner products lies within a special class of matrices.

## 4.1 The Dot Product's "Engine": The Identity Matrix

We can express the standard dot product using the identity matrix  $I$ :

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T I \mathbf{y}$$

The identity matrix  $I$  is the "engine" driving the dot product. To build a new inner product, we just swap out this engine.

### Example 1. The Dot Product as a Matrix Operation

Let's verify this for  $\mathbf{x} = [1, 2]^T$  and  $\mathbf{y} = [3, 4]^T$ .

1. **Direct Calculation:**  $\mathbf{x} \cdot \mathbf{y} = (1)(3) + (2)(4) = 3 + 8 = 11$ .
2. **Matrix Calculation:**

$$\mathbf{x}^T I \mathbf{y} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = 11$$

The results match perfectly. This confirms that the identity matrix  $I$  is the engine for the standard dot product.

## 4.2 Custom Inner Products with New Matrices

### Theorem 1. Inner Product Matrix Condition

An operation defined by  $\langle \mathbf{x}, \mathbf{y} \rangle_A = \mathbf{x}^T A \mathbf{y}$  is a valid inner product if and only if the matrix  $A$  is **symmetric** and **positive-definite**.

This theorem is our master recipe. By choosing a different symmetric, positive-definite matrix  $A$ , we can define a completely new geometry.

### Example 2. Verifying a Custom Engine

Let's verify that  $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix}$  defines a valid inner product.

**1. Is A Symmetric?** Yes. The matrix is unchanged if we flip it across its main diagonal.

**2. Is A Positive-Definite?** We must prove that for any non-zero vector  $\mathbf{x} = [a, b, c]^T$ , the value of  $\mathbf{x}^T A \mathbf{x}$  is always positive. We found this to be:

$$\mathbf{x}^T A \mathbf{x} = 2a^2 + 2ab + 3b^2 - 2bc + 2c^2$$

By "completing the square," we can rewrite this as:

$$(a+b)^2 + (b-c)^2 + a^2 + b^2 + c^2$$

This is a sum of five squared terms. For any non-zero vector, at least one of  $a, b, c$  must be non-zero, guaranteeing the sum is positive.

## 4.3 Why Positive-Definiteness is Essential: A Counterexample

What happens if we choose a matrix that is symmetric but *not* positive-definite? Let's investigate with a simple example.

Consider the symmetric matrix  $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . This matrix is not positive-definite. Let's see why. Let's pick a vector  $\mathbf{x} = [0, 1]^T$ . Now, let's compute the inner product of this vector with itself, which should give us its squared length.

$$\langle \mathbf{x}, \mathbf{x} \rangle_A = \mathbf{x}^T A \mathbf{x} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -1$$

The "squared length" of our non-zero vector is -1. If we try to find its length (or norm), we get:

$$\|\mathbf{x}\|_A = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_A} = \sqrt{-1} = i$$

### Q Geometric Breakdown

The length of a non-zero vector is an imaginary number! This is a complete breakdown of our geometric intuition. A physical length cannot be imaginary, and its square certainly cannot be negative.

This is why the positive-definiteness axiom is non-negotiable. It is the rule that ensures our generalized concept of length behaves like a real, physical length. It guarantees that the length of any vector is a real, non-negative number.

## 4.4 The Geometric Interpretation: How Matrix A Reshapes Space

What is the matrix  $A$  in  $\langle \mathbf{x}, \mathbf{y} \rangle_A = \mathbf{x}^T A \mathbf{y}$  actually doing? It is geometrically reshaping or transforming the vector space.

Let's think about all the vectors that have a length of 1. In standard Euclidean space (where  $A = I$ ), the set of all vectors  $\mathbf{x} = [x, y]^T$  with norm 1 satisfies  $\|\mathbf{x}\| = \sqrt{x^2 + y^2} = 1$ , which is the equation of the familiar **unit circle**.

Now, let's impose a new inner product with the matrix  $A = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$ . This matrix is symmetric and positive-definite. What does the set of "unit vectors" look like in *this*

new geometry? We need to find all vectors  $\mathbf{x}$  such that their new norm is 1.

$$\|\mathbf{x}\|_A = \sqrt{\mathbf{x}^T A \mathbf{x}} = \sqrt{4x^2 + y^2} = 1$$

Squaring both sides gives us the equation:

$$4x^2 + y^2 = 1 \quad \text{or} \quad \frac{x^2}{(1/2)^2} + \frac{y^2}{1^2} = 1$$

This is not the equation of a circle. It's the equation of an **ellipse** that is twice as tall as it is wide.

### Q From Unit Circle to Unit Ellipse

The inner product matrix  $A$  has transformed our space. What used to be the unit circle is now a unit ellipse.

This means our perception of distance has changed. In this new space, the vector  $[1/2, 0]^T$  is now considered a "unit vector" because its new length is 1. We have effectively "squashed" the space horizontally. Distances along the x-axis are magnified compared to distances along the y-axis.

If we had used a non-positive-definite matrix (like the one in our counterexample), the shape of "unit vectors" would not be a closed ellipse but an open shape like a hyperbola. This doesn't provide a coherent sense of "near" and "far," which again shows why positive-definiteness is essential for creating sensible geometries.

# 5 What Our New Tools Can Do

---

Every inner product we create comes with a complete geometric toolkit.

## 5.1 Custom Lengths and Distances

An inner product first gives us a specific way to measure length, called a **norm**.

### Definition 3. Norm

The **norm** of a vector  $\mathbf{x}$  induced by an inner product is:

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

From the norm, we can define the **distance** (or metric) between two vectors as the length of the vector connecting them.

### Definition 4. Distance (Metric)

The **distance** (or metric) between vectors  $\mathbf{x}$  and  $\mathbf{y}$  is:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle}$$

## 5.2 Redefining Angles and Orthogonality

The inner product also defines the angle  $\omega$  between two vectors, which is a perfect measure of their similarity.

$$\cos(\omega) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$$

### Definition 5. Orthogonality

Two vectors  $\mathbf{x}$  and  $\mathbf{y}$  are **orthogonal** with respect to an inner product if  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ .

#### 💡 Key Takeaway

**Orthogonality is relative.** Two vectors might be perfectly perpendicular in one geometry but not in another.

**Example 3. Orthogonality is in the Eye of the Beholder**

Consider  $\mathbf{x} = [1, 1]^T$  and  $\mathbf{y} = [-1, 1]^T$ .

**Case 1: Geometry of the Dot Product**

$$\langle \mathbf{x}, \mathbf{y} \rangle = (1)(-1) + (1)(1) = 0$$

The vectors are orthogonal.

**Case 2: Geometry defined by  $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$**

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y} \rangle_A &= \mathbf{x}^T A \mathbf{y} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ &= \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = -1\end{aligned}$$

The inner product is not zero, so the vectors are *not* orthogonal in this new geometry.

## 6 Orthonormal Bases

---

Working in any geometric space is much easier if you have a good coordinate system. In vector spaces, the ultimate "grid" is an **orthonormal basis**.

### Definition 6. Orthonormal Basis

An **orthonormal basis** is a set of basis vectors  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  where every vector has a norm of 1, and every vector is orthogonal to all the others.

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

If we have a valid basis, but it's "messy"—the vectors aren't unit length or orthogonal—we can clean it up using the **Gram-Schmidt Orthonormalization Process**.

#### Q The Gram-Schmidt Idea: Subtracting the "Shadows"

The process works by building the orthonormal basis one vector at a time by removing any components that are not orthogonal.

#### Example 4. A Narrated Gram-Schmidt Calculation

Let's build an orthonormal basis from the basis vectors  $\mathbf{x}_1 = [1, 1, 0]^T$  and  $\mathbf{x}_2 = [1, 2, 2]^T$ . We will use the standard dot product.

**Step 1: Process  $\mathbf{x}_1$  to find the first clean vector  $\mathbf{e}_1$**  First, we normalize  $\mathbf{x}_1$  by dividing by its length.  $\|\mathbf{x}_1\| = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2}$ .

$$\mathbf{e}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \frac{1}{\sqrt{2}}[1, 1, 0]^T = \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right]^T$$

**Step 2: Process  $\mathbf{x}_2$  to find the second clean vector  $\mathbf{e}_2$**  First, we create a new vector,  $\mathbf{u}_2$ , that is orthogonal to  $\mathbf{e}_1$  by subtracting the projection of  $\mathbf{x}_2$  on  $\mathbf{e}_1$ . The projection (shadow) is  $\text{proj}_{\mathbf{e}_1} \mathbf{x}_2 = \langle \mathbf{x}_2, \mathbf{e}_1 \rangle \mathbf{e}_1$ .

$$\langle \mathbf{x}_2, \mathbf{e}_1 \rangle = (1)\frac{1}{\sqrt{2}} + (2)\frac{1}{\sqrt{2}} + (2)(0) = \frac{3}{\sqrt{2}}$$

Now, subtract the shadow from  $\mathbf{x}_2$ :

$$\begin{aligned}\mathbf{u}_2 &= \mathbf{x}_2 - \frac{3}{\sqrt{2}}\mathbf{e}_1 \\ &= [1, 2, 2]^T - \frac{3}{\sqrt{2}} \left[ \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right]^T \\ &= [1, 2, 2]^T - \left[ \frac{3}{2}, \frac{3}{2}, 0 \right]^T = \left[ -\frac{1}{2}, \frac{1}{2}, 2 \right]^T\end{aligned}$$

Now we normalize  $\mathbf{u}_2$ .  $\|\mathbf{u}_2\| = \sqrt{(-\frac{1}{2})^2 + (\frac{1}{2})^2 + 2^2} = \sqrt{\frac{1}{4} + \frac{1}{4} + 4} = \sqrt{\frac{18}{4}} = \frac{3\sqrt{2}}{2}$ .

$$\mathbf{e}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{2}{3\sqrt{2}} \left[ -\frac{1}{2}, \frac{1}{2}, 2 \right]^T = \left[ -\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3} \right]^T$$

**Result:** Our new, clean orthonormal basis is  $\{\mathbf{e}_1, \mathbf{e}_2\}$ .

## 7 Conclusion: A New Perspective on Geometry

---

Our investigation started with a practical need: to measure similarity in a way that makes sense for a specific problem. This led us on a conceptual journey away from the familiar comfort of Euclidean distance and into a world of custom geometries.

We began by taking apart our standard ruler, the dot product, and identifying the three fundamental principles that gave it power: linearity, symmetry, and positive-definiteness. By treating these principles not as observations but as a blueprint, we discovered the general inner product—a tool that lets us define new ways of measuring.

The master recipe for this new tool, we learned, is the symmetric, positive-definite matrix. This matrix is more than just a collection of numbers; it is the engine that actively reshapes vector space, transforming our very notions of length, distance, and what it means to be perpendicular. We saw that without the crucial property of positive-definiteness, our sense of geometry collapses into contradictions like imaginary lengths.

This journey teaches us a profound lesson: geometry is not absolute. It is a lens through which we view data, and we have the power to craft the lens that best suits our problem. The inner product provides the framework, the matrix provides the specific design, and methods like Gram-Schmidt ensure we can always build a simple, orderly grid to navigate any geometric world we create.