Inner Products

Definition 15.1 (Inner Product): The Euclidean inner product between two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ is defined as:

$$\langle \vec{x}, \vec{y} \rangle \equiv \vec{x}^T \vec{y} \tag{1}$$

$$= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{2}$$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \tag{3}$$

$$= \sum_{i=1}^{n} x_i y_i \tag{4}$$

In physics, inner products are often called dot products (denoted by $\vec{x} \cdot \vec{y}$) and can be used to calculate values such as work. In this class we will use $\langle \vec{x}, \vec{y} \rangle$ for the inner product.

Example 15.1 (Inner product of two vectors): Compute the inner product of the two vectors $\begin{bmatrix} -1 & 3.5 & 0 \end{bmatrix}^T$ and $\begin{bmatrix} 1 & 0 & 2 \end{bmatrix}^T$.

$$\left\langle \begin{bmatrix} -1 \\ 3.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\rangle = \begin{bmatrix} -1 & 3.5 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \tag{5}$$

$$= -1 \times 1 + 3.5 \times 0 + 0 \times 2 \tag{6}$$

$$= -1 + 0 + 0 = -1. \tag{7}$$

Properties of Inner Products

Symmetry We can prove algebraically that the inner product is a symmetric operation, that is, it remains the same even if we switch its arguments:

$$\begin{aligned} \langle \vec{x}, \vec{y} \rangle &= x_1 y_1 + \cdots + x_n y_n \\ &= y_1 x_1 + \cdots + y_n x_n \\ &= \langle \vec{y}, \vec{x} \rangle \end{aligned} \tag{8}$$

Homogeneity We shall see how scaling affects inner products. If c is a real number, algebraically:

$$\begin{aligned} \langle c\vec{x}, \vec{y} \rangle &= (cx_1)y_1 + \cdots + (cx_n)y_n \\ &= c(x_1y_1) + \cdots + c(x_ny_n) \\ &= c\langle \vec{x}, \vec{y} \rangle \end{aligned} \tag{9}$$

By a similar argument (exercise: prove it yourself),

$$\langle \vec{x}, c\vec{y} \rangle = c\langle \vec{x}, \vec{y} \rangle \tag{10}$$

This means the scaling any vector in the inner product by any real constant c will scale the inner product by the same constant.

Additivity What happens when we take the inner product between a sum of two vectors and another vector? We can write:

$$\begin{aligned} \langle \vec{x} + \vec{y}, \vec{z} \rangle &= (x_1 + y_1)z_1 + \cdots + (x_n + y_n)z_n \\ &= x_1z_1 + y_1z_1 + \cdots + x_nz_n + y_nz_n \\ &= x_1z_1 + \cdots + x_nz_n + y_1z_1 + \cdots + y_nz_n \\ &= \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle \end{aligned} \tag{11}$$

Similarily:

$$\langle \vec{x}, \vec{y} + \vec{z} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle \tag{12}$$

Homogeneity together with additivity gives rise to a property called *Bilinearity*. This means that the inner product is a function that is linear in each argument. These properties are also very useful in proving properties about inner products.

Orthogonal Vectors

Two vectors $\vec{x}$, $\vec{y}$ are said to be **orthogonal** if $\langle \vec{x}, \vec{y} \rangle = 0$. In 2-D and 3-D coordinate spaces, perpendicular and orthogonal mean the same thing; in higher-dimension spaces, it is harder to visualize vectors being perpendicular, so the term orthogonal comes in handy to abstract away the need to visualize.

Here's an example, in $\mathbb{R}^3$. Let's say $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} 4 \\ -1 \\ -1 \end{bmatrix}$. Then, $\langle \vec{x}, \vec{y} \rangle = (1)(4) + (1)(-1) + (3)(1) = 0$. Thus $\vec{x}$ and $\vec{y}$ are orthogonal.

A very important fact is that the standard unit vectors are always orthogonal to each other.

Special Vector Operations

The inner product is a basic building block for many operations. In programming terms, if you have a black box function INNERPRODUCT, here are some useful operations you can form by controlling what vectors you input into the function.

Sum of Components

$$\left\langle \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T, \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix}^T \right\rangle = x_0 + x_1 + \ldots + x_n \tag{13}$$

Average

$$\left\langle \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \ldots & \frac{1}{n} \end{bmatrix}^T, \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix}^T \right\rangle = \frac{x_0 + x_1 + \ldots + x_n}{n} \tag{14}$$

Sum of Squares

$$\left\langle \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix}^T, \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix}^T \right\rangle = x_0^2 + x_1^2 + \ldots + x_n^2 \tag{15}$$

Selective Sum Here, the values of the first vector has a set of 0's and 1's, which correspond to which values of the variable matrix (the vertical one) are to be used, and which ones are to be thrown away for the sum. This, for example, can take the place of a for loop checking every element of a vector.

$$\left\langle \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & \cdots & 0 & 1 \end{bmatrix}^T, \begin{bmatrix} x_0 & x_1 & \cdots & x_n & \end{bmatrix}^T \right\rangle = x_2 + x_4 + \ldots + x_n \tag{16}$$

These things are often important in computer programming contexts because computers (and programming languages) are often optimized to be able to do vector operations like inner products very fast. So, seeing an inner-product way of representing something can often speed up calculations considerably.

Introduction to Norms

The **Euclidean Norm** of a vector is given by:

$$\|\vec{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \sqrt{\langle \vec{x}, \vec{x} \rangle} \tag{17}$$

This corresponds to the usual notion of distance in $\mathbb{R}^2$ or $\mathbb{R}^3$. It is interesting to note that the set of points with equal Euclidean norm is a circle in $\mathbb{R}^2$, or a sphere in $\mathbb{R}^3$.

You may have noticed that the subscript 2 in the definition of the norm given above. The subscript differentiates the Euclidean norm (or 2-norm) from other useful norms. In general, the p-norm is defined as:

$$\|\vec{x}\|_p = (x_1^p + x_2^p + \ldots + x_n^p)^{1/p} \tag{18}$$

These other norms might feel esoteric but turn out to be spectacularly useful in many engineering settings. Follow-on courses like 127 will show you how these can be useful in applications like machine learning.

Why is the norm important? The 2-norm of a vector is also the magnitude of the vector (or length of the arrow).

Properties of Norms

$$\|\alpha x\| = |\alpha| \, \|x\| \tag{19}$$

$$\|x + y\| \leq \|x\| + \|y\| \text{ known as the "triangle inequality"} \tag{20}$$

$$\|x\| \geq 0 \tag{21}$$

$$\|x\| = 0 \text{ only if } x = 0 \tag{22}$$

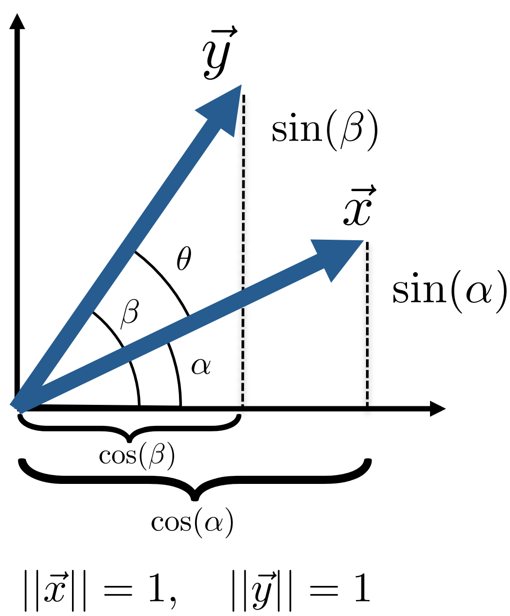
Interpretations of Inner Products

Now that we have defined inner products, it's time to get an intuition about what an inner product is.

First we take the unit vector $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and a general unit vector in $\mathbb{R}^2$, $\vec{x} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$. $\vec{x}$ is a unit vector because $\|\vec{x}\| = \sqrt{\cos^2\alpha + \sin^2\alpha} = 1$.

When we draw the vectors, the angle between them is clearly α. We can calculate $\langle \vec{e}_1, \vec{y} \rangle = 1 \times \cos\alpha + 0 \times \sin\alpha = \cos\alpha$. With this we might guess that the inner product between two unit vectors is the cosine of the angle between them.

Let's try to prove this for two general unit vectors in $\mathbb{R}^2$, $\vec{x} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix}$.



Using the trigonometric identity $\cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha - \beta)$, we find:

$$\begin{aligned} \langle \vec{x}, \vec{y} \rangle &= \cos\alpha\cos\beta + \sin\alpha\sin\beta \\ &= \cos(\alpha - \beta) \\ &= \cos\theta \end{aligned} \tag{23}$$

Again, that is the cosine of the angle between these two vectors!

What about any two general vectors $\vec{x}$ and $\vec{y}$? We can first convert them into the unit vectors $\vec{x}/\|x\|$ and $\vec{y}/\|y\|$ by dividing each vector by its norm. Then we can use (23) to get:

$$\left\langle \frac{\vec{x}}{\|x\|}, \frac{\vec{y}}{\|y\|} \right\rangle = \cos\theta \tag{24}$$

We can then use the homogeneity property (9) (since both norms are scalars):

$$\begin{aligned}\frac{\langle \vec{x}, \vec{y} \rangle}{||x||\,||y||} &= \cos\theta \\ \langle \vec{x}, \vec{y} \rangle &= ||x||\,||y||\cos\theta\end{aligned} \tag{25}$$

Now we have a geometric interpretation of the inner product: the inner product of vectors $\vec{x}$ and $\vec{y}$ is their lengths multiplied by the angle between them. One remarkable observation is that the inner product *does not depend on the coordinate system the vectors are in*, it only depends on the relative angle between these vectors and their length. This is the reason it is very useful in physics, where the physical laws do not depend on the coordinate system used to measure them. This is also the reason why this property holds in higher dimensions as well. For any two vectors we can look at the plane passing through them and the angle between them is the angle θ measured in the plane.

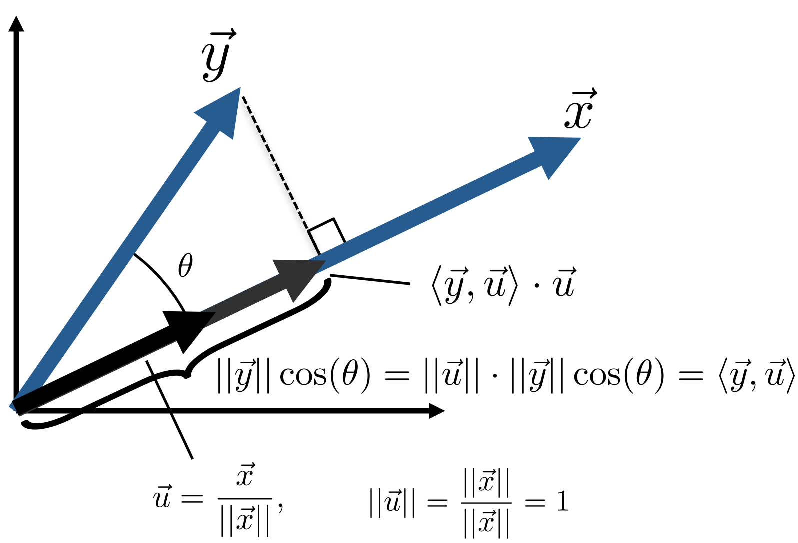
The Cauchy-Schwarz Inequality

Now we are ready to prove the Cauchy-Schwarz inequality, which relates the inner product of two vectors to their length. Since $|\cos\theta| \leq 1$ for all real θ,

$$\begin{aligned}|\langle \vec{x}, \vec{y} \rangle| &= |\,||\vec{x}||\,||\vec{y}||\cos\theta| \\ &= ||\vec{x}||\,||\vec{y}|| \cdot |\cos\theta| \\ &\leq ||\vec{x}||\,||\vec{y}||\end{aligned} \tag{26}$$

Projections

Knowing that the inner products of two vectors in $\mathbb{R}^n$ is their lengths multiplied by the angle between them, we can write the *projection* of a vector onto another using the inner product.



The projection of $\vec{y}$ onto $\vec{x}$ is the component of $\vec{y}$ lying in the direction of $\vec{x}$. First let's find the length of this component. By simple geometry, this is $||\vec{y}||\cos\theta$, where θ is the angle between $\vec{y}$ and $\vec{x}$. We can write it as

$\langle \vec{y}, \vec{u} \rangle$, where $\vec{u} = \vec{x} / \|\vec{x}\|$ is the unit vector in the direction of $\vec{x}$. Now we know the length of the projection, all we need is a direction. This is just $\vec{u}$, so the projection of $\vec{x}$ onto $\vec{y}$ is given by:

$$\langle \vec{y}, \vec{u} \rangle \vec{u} = \left\langle \vec{y}, \frac{\vec{x}}{\|\vec{x}\|} \right\rangle \frac{\vec{x}}{\|\vec{x}\|} \tag{27}$$