Dot product

The **dot product** is an inner product on a coordinate vector space (Definition 1, Theorem 1).

Definition 1 *Given vectors* $\mathbf{v}$ *and* $\mathbf{u}$ *in n-dimensional space, the dot product is defined as,*

$$\mathbf{v} \cdot \mathbf{u} = \sum_{i=1}^{n} v_i u_i$$

Theorem 1 *The axioms of an inner product hold for the dot product. That is:*

1. $(\mathbf{v} + \mathbf{u}) \cdot \mathbf{w} = (\mathbf{v} \cdot \mathbf{w}) + (\mathbf{u} \cdot \mathbf{w})$
2. $(\alpha \mathbf{v} \cdot \mathbf{u}) = \alpha(\mathbf{v} \cdot \mathbf{u})$
3. $\mathbf{v} \cdot \mathbf{u} = \mathbf{u} \cdot \mathbf{v}$
4. $\mathbf{v} \cdot \mathbf{v} \geq 0$ *and* $\mathbf{v} \cdot \mathbf{v} = 0 \iff \mathbf{v} = \mathbf{0}$

Proof:

1.)

$$\begin{aligned}
(\mathbf{v} + \mathbf{u}) \cdot \mathbf{w} &= \sum_{i=1}^{d} (v_i + u_i) w_i \\
&= \sum_{i=1}^{d} (v_i w_i + u_i w_i) \\
&= \sum_{i=1}^{d} v_i w_i + \sum_{i=1}^{d} u_i w_i \\
&= (\mathbf{v} \cdot \mathbf{w}) + (\mathbf{u} \cdot \mathbf{w})
\end{aligned}$$

2.)

$$\begin{aligned}(\alpha\mathbf{v}\cdot\mathbf{u}) &= \sum_{i=1}^{d}(\alpha v_i)w_i \\ &= \alpha\sum_{i=1}^{d}(v_i w_i) \\ &= \alpha(\mathbf{v}\cdot\mathbf{u})\end{aligned}$$

3.)

$$\begin{aligned}\mathbf{v}\cdot\mathbf{u} &= \sum_{i=1}^{d} v_i u_i \\ &= \sum_{i=1}^{d} u_i v_i \\ &= \mathbf{u}\cdot\mathbf{v}\end{aligned}$$

4.)

$$\mathbf{v}\cdot\mathbf{v} = \sum_{i=1}^{d} v_i^2$$

and

$$\sum_{i=1}^{d} v_i^2 = 0 \iff \forall i \in [d]\ v_i = 0 \iff \mathbf{v} = \mathbf{0}$$

and

$$\exists i \in [d] \text{ s.t. } v_i \neq 0 \implies \sum_{i=1}^{d} v_i^2 > 0$$

□

Intuition

There are three perspectives I find useful for thinking about the dot products. Ordered from the least abstract to the most abstract, these perspectives are:

1. *The dot product succinctly describes a weighted sum*

2. *The dot product describes a geometric relationship between two vectors*

3. *The dot product is analogous to the product on scalars*

These perspectives are described in the sections below:

1. The dot product succinctly describes a weighted sum

The dot product is useful for succinctly describing a weighted sum of variables. Lets say we have a vector of variables storing some kind of data $\mathbf{x}$. Let's say we have a vector of weights $\mathbf{w}$ and we want to sum the variables in $\mathbf{x}$ where each element x_i in $\mathbf{x}$ is multiplied by its weight w_i in $\mathbf{w}$. This operation is stated succinctly as $\mathbf{w} \cdot \mathbf{x}$. For example, if the weight vector $\mathbf{w}$ consists of only ones, then the dot product is simply the sum of the variables in $\mathbf{x}$.

Whenever you find a dot product, it helps to think about the operation as a sum of variables where each variable is multiplied by a weight.

2. The dot product describes a geometric relationship between two vectors

The dot product uses the relationship between the directions in which the two vectors point. More specifically, if the two vectors point in a similar direction, the magnitude of the dot product increases. If they point in drastically different directions, the dot product decreases. Now, the question naturally arises, what do we mean by "point in a similar direction?" What do we mean by "similar"? The dot product asserts that the angle between the two vectors measures how similarly they point. The smaller the angle, the smaller will be the dot product. To show how this works, we first show that the dot product can be computed using the angle between $\mathbf{a}$ and $\mathbf{b}$ as follows:

$$\mathbf{a} \cdot \mathbf{b} = ||\mathbf{a}|| \, ||\mathbf{b}|| \cos \theta$$

(Theorem 2). If $\theta := 0$, then the two vectors point in the same direction. In this case, $\cos \theta = 1$ and the dot product reduces to simply computing the product of the two vectors' magnitudes. If $\theta = \pi/2$, then the two vectors point in perpendicular directions (i.e. maximally different directions). We see that $\cos \pi/2 = 0$ and the dot product between the two vectors is zero.

Another way to understand how this works is to look at the projection of one vector onto the other. That is, given two vectors $\mathbf{a}$, $\mathbf{b}$, the dot product between these vectors computes the product of the magnitudes of $\mathbf{a}$ and $\mathbf{b}$ along the direction that the two vectors share (Figure 1). Said differently, the dot product $\mathbf{a} \cdot \mathbf{b}$ can be viewed as the magnitude of the projection of one of the vectors onto the other vector multiplied by the magnitude of the vector being projected upon. That is,

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= \left|\left|\text{proj}(\mathbf{a}, \mathbf{b})\right|\right| \, ||\mathbf{b}|| \\ &= \left|\left|\text{proj}(\mathbf{b}, \mathbf{a})\right|\right| \, ||\mathbf{a}|| \end{aligned}$$

If the two vectors are orthogonal, then the projection of either vector onto the other will be zero and thus the dot product will be zero. In contrast, if two vectors point in the same direction, then the projection of the smaller vector onto the larger vector is simply the smaller vector so we multiply the magnitude of the smaller vector by the magnitude of the larger vector (i.e. simply multiply their norms).

Given this geometric interpretation of the dot product, we can see that taking the dot product of some vector **a** and a *unit vector* **b**, finds the length of the projection of **a** along the axis defined by **b**:

$$\begin{aligned}\mathbf{a} \cdot \mathbf{b} &= ||\mathbf{a}|| \left\|\text{proj}(\mathbf{b}, \mathbf{a})\right\| \\ &= \left\|\text{proj}(\mathbf{b}, \mathbf{a})\right\| && \text{because } ||\mathbf{a}|| = 1\end{aligned}$$

Thus, whenever one of the vectors in a dot product is a unit vector, the operation can always be viewed as the length of the projection along the axis defined by the unit vector.

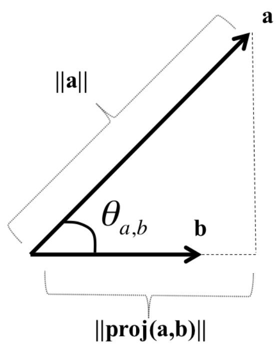

Figure 1: The projection of **a** onto **b**.

Theorem 2 *Given θ is the angle between the two* **v** *and* **w**, *the following definition for the dot product between* **v** *and* **w** *is equivalent to Definition 1:*

$$\mathbf{v} \cdot \mathbf{w} = ||\mathbf{v}|| \, ||\mathbf{w}|| \cos\theta$$

Proof:

Let **a** and **b** be two vectors in a d-dimensional coordinate space and let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d$

be the standard basis vectors of the space. Then,

$$\mathbf{a} = \sum_{i=1}^{d} a_i \mathbf{e}_i$$

$$\mathbf{b} = \sum_{i=1}^{d} b_i \mathbf{e}_i$$

First, we note that by the definition of a standard basis, the vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d$ are all orthonormal to each other. That is,

$$\begin{aligned} \mathbf{e}_i \cdot \mathbf{e}_i &= 1 && \text{they are all unit vectors} \\ i \neq j \implies \mathbf{e}_i \cdot \mathbf{e}_j &= 0 && \text{they are all orthogonal to each other} \end{aligned}$$

Next, by the geometric definition of the dot product,

$$\begin{aligned} \mathbf{a} \cdot \mathbf{e}_i &= ||\mathbf{a}|| \, ||\mathbf{e}_i|| \cos \theta_{\mathbf{a},\mathbf{e}_i} \\ &= ||\mathbf{a}|| \cos \theta_{\mathbf{a},\mathbf{e}_i} \\ &= a_i && \text{see Figure 1} \end{aligned}$$

We see that a_i is the component of $\mathbf{a}$ in the direction of the base-vector $\mathbf{e}_i$. Finally,

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= \mathbf{a} \cdot \left(\sum_{i=1}^{d} b_i \mathbf{e}_i \right) \\ &= \sum_{i=1}^{d} (\mathbf{a} \cdot b_i \mathbf{e}_i) && \text{axiom 1 of inner product} \\ &= \sum_{i=1}^{d} b_i (\mathbf{e}_i \cdot \mathbf{a}) && \text{axiom 3 of inner product} \\ &= \sum_{i=1}^{d} b_i (\mathbf{a} \cdot \mathbf{e}_i) && \text{axiom 2 of inner product} \\ &= \sum_{i=1}^{d} b_i a_i \end{aligned}$$

This is the algebraic definition of the dot product.

□

3. The dot product is analogous to the product on scalars

One way to understand the dot product is as an operation on vectors that is analogous to the product on scalars. Given two scalars, $x, y \in \mathbb{R}$, it is obvious that the more we increase the magnitude (i.e. absolute value) of either x or y, the more that the magnitude of the product will grow. A product on vectors should behave similarly and indeed the dot product does. Given two vectors **a** and **b**, if we increase the norm of either of the vectors, the magnitude of the dot product increases. We see this clearly expressed in the $||\mathbf{a}|| \, ||\mathbf{b}||$ term of the geometric definition of the dot product:

$$\mathbf{a} \cdot \mathbf{b} = ||\mathbf{a}|| \, ||\mathbf{b}|| \cos \theta$$

However, unlike the product between scalars, a product between vectors should also take into account the direction in which the two vectors point. The dot product asserts that if the two vectors point in a similar direction, the magnitude of the dot product increases. If they point in drastically different directions, the dot product decreases.

Antoerh aspect to the product of scalars that is analogous to the dot product is that if x and y have opposite signs then $xy < 0$. Can the idea of "opposite signs" be expressed in a product on vectors? Since the norm is always positive, the term, $||\mathbf{a}|| \, ||\mathbf{b}||$, cannot express "opposite signs." Rather, we see that if the angle between the vectors **a** and **b** is obtuse, then their dot product will be negative:

$$\begin{aligned} -\frac{\pi}{2} > \theta_{\mathbf{a},\mathbf{b}} > -\frac{3\pi}{2} &\implies \cos \theta_{\mathbf{a},\mathbf{b}} < 0 \\ &\implies ||\mathbf{a}|| \, ||\mathbf{b}|| \cos \theta_{\mathbf{a},\mathbf{b}} < 0 \\ &\implies \mathbf{a} \cdot \mathbf{b} < 0 \end{aligned}$$

Thus, two vectors "have opposite signs", according to the dot product, if the angle between them is greater than $\pi/2$ and less than $3\pi/4$.