Lecture 3: The Dot Product

3.1 The angle between vectors

Suppose $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ are two vectors in $\mathbb{R}^2$, neither of which is the zero vector $\mathbf{0}$. Let α and β be the angles between $\mathbf{x}$ and $\mathbf{y}$ and the positive horizontal axis, respectively, measured in the counterclockwise direction. Supposing $\alpha \geq \beta$, let $\theta = \alpha - \beta$. Then θ is the angle between $\mathbf{x}$ and $\mathbf{y}$ measured in the counterclockwise direction. From the subtraction formula for cosine we have

$$\cos(\theta) = \cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta).$$

Now

$$\cos(\alpha) = \frac{x_1}{|\mathbf{x}|},$$

$$\cos(\beta) = \frac{y_1}{|\mathbf{y}|},$$

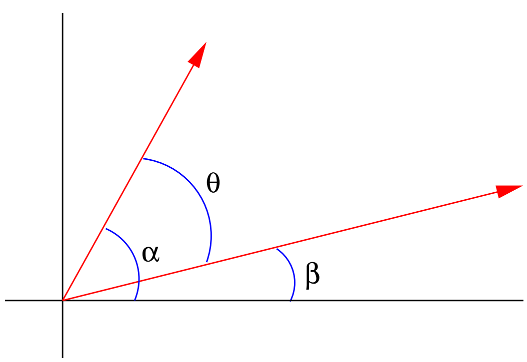
$$\sin(\alpha) = \frac{x_2}{|\mathbf{x}|},$$

and

$$\sin(\beta) = \frac{y_2}{|\mathbf{y}|}.$$

Thus, we have

$$\cos(\theta) = \frac{x_1 y_1}{|\mathbf{x}||\mathbf{y}|} + \frac{x_2 y_2}{|\mathbf{x}||\mathbf{y}|} = \frac{x_1 y_1 + x_2 y_2}{|\mathbf{x}||\mathbf{y}|}.$$



The angle between two vectors in $\mathbb{R}^2$

Example Let θ be the smallest angle between $\mathbf{x} = (2, 1)$ and $\mathbf{y} = (1, 3)$, measured in the counterclockwise direction. Then

$$\cos(\theta) = \frac{(2)(1) + (1)(3)}{|\mathbf{x}||\mathbf{y}|} = \frac{5}{\sqrt{5}\sqrt{10}} = \frac{1}{\sqrt{2}}.$$

Hence

$$\theta = \cos^{-1}\left(\frac{1}{\sqrt{2}}\right) = \frac{\pi}{4}.$$

With more work it is possible to show that if $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$ are two vectors in $\mathbb{R}^3$, neither of which is the zero vector $\mathbf{0}$, and θ is the smallest positive angle between $\mathbf{x}$ and $\mathbf{y}$, then

$$\cos(\theta) = \frac{x_1y_1 + x_2y_2 + x_3y_3}{|\mathbf{x}||\mathbf{y}|}$$

3.2 The dot product

Definition If $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ are vectors in $\mathbb{R}^n$, then the *dot product* of $\mathbf{x}$ and $\mathbf{y}$, denoted $\mathbf{x} \cdot \mathbf{y}$, is given by

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Note that the dot product of two vectors is a scalar, not another vector. Because of this, the dot product is also called the *scalar product.* It is also an example of what is called an *inner product* and is often denoted by $\langle \mathbf{x}, \mathbf{y} \rangle$.

Example If $\mathbf{x} = (1, 2, -3, -2)$ and $\mathbf{y} = (-1, 2, 3, 5)$, then

$$\mathbf{x} \cdot \mathbf{y} = (1)(-1) + (2)(2) + (-3)(3) + (-2)(5) = -1 + 4 - 9 - 10 = -16.$$

Proposition For any vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ in $\mathbb{R}^n$ and scalar α,

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

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

$$(\alpha\mathbf{x}) \cdot \mathbf{y} = \alpha(\mathbf{x} \cdot \mathbf{y}),$$

$$\mathbf{0} \cdot \mathbf{x} = 0,$$

$$\mathbf{x} \cdot \mathbf{x} \geq 0,$$

$$\mathbf{x} \cdot \mathbf{x} = 0 \text{ only if } \mathbf{x} = \mathbf{0},$$

and

$$\mathbf{x} \cdot \mathbf{x} = |\mathbf{x}|^2.$$

At this point we can say that if $\mathbf{x}$ and $\mathbf{y}$ are two nonzero vectors in either $\mathbb{R}^2$ or $\mathbb{R}^3$ and θ is the smallest positive angle between $\mathbf{x}$ and $\mathbf{y}$, then

$$\cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|}.$$

We would like to be able to make the same statement about the angle between two vectors in any dimension, but we would first have to define what we mean by the angle between two vectors in $\mathbb{R}^n$ for $n > 3$. The simplest way to do this is to turn things around and use the dot product to define the angle. However, in order for this to work we must first know that

$$-1 \leq \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|} \leq 1,$$

since this is the range of values for the cosine function. This fact follows from the following inequality.

Cauchy-Schwarz Inequality For all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$,

$$|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|.$$

Proof To see why this is so, first note that both sides of of the inequality are 0 when $y = \mathbf{0}$, and hence are equal in this case. Assuming $\mathbf{x}$ and $\mathbf{y}$ are fixed vectors in $\mathbb{R}^n$, with $\mathbf{y} \neq \mathbf{0}$, let t be a real number and consider the function

$$f(t) = (\mathbf{x} + t\mathbf{y}) \cdot (\mathbf{x} + t\mathbf{y}).$$

Now $f(t) \geq 0$ for all t and, moreover,

$$f(t) = \mathbf{x} \cdot \mathbf{x} + \mathbf{x} \cdot t\mathbf{y} + t\mathbf{y} \cdot \mathbf{x} + t\mathbf{y} \cdot t\mathbf{y} = |\mathbf{x}|^2 + 2(\mathbf{x} \cdot \mathbf{y})t + |\mathbf{y}|^2 t^2.$$

Hence f is a quadratic polynomial with at most one root. Since the roots of f are, as given by the quadratic formula,

$$\frac{-2(\mathbf{x} \cdot \mathbf{y}) \pm \sqrt{4(\mathbf{x} \cdot \mathbf{y})^2 - 4|\mathbf{x}|^2|\mathbf{y}|^2}}{2|\mathbf{y}|^2},$$

it follows that we must have

$$4(\mathbf{x} \cdot \mathbf{y})^2 - 4|\mathbf{x}|^2|\mathbf{y}|^2 \leq 0.$$

Thus

$$(\mathbf{x} \cdot \mathbf{y})^2 \leq |\mathbf{x}|^2|\mathbf{y}|^2,$$

and so

$$|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|.$$

Note that $|\mathbf{x} \cdot \mathbf{y}| = |\mathbf{x}||\mathbf{y}|$ if and only if there is some value of t for which $f(t) = 0$, which happens if and only if $\mathbf{x} + t\mathbf{y} = \mathbf{0}$, that is, $\mathbf{x} = -t\mathbf{y}$, for some value of t. Moreover, if $\mathbf{y} = \mathbf{0}$, then $\mathbf{0} = 0\mathbf{x}$ for any $\mathbf{x}$ in $\mathbb{R}^n$. Hence, in either case, the Cauchy-Schwarz inequality becomes an equality if and only if either $\mathbf{x}$ is a scalar multiple of $\mathbf{y}$ or $\mathbf{y}$ is a scalar multiple of $\mathbf{x}$. If $\mathbf{x}$ and $\mathbf{y}$ are nonzero vectors, then we have equality if and only if $\mathbf{x}$ and $\mathbf{y}$ are parallel.

With the Cauchy-Schwarz inequality we have

$$-1 \leq \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|} \leq 1,$$

for any nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. Thus we may now state the following definition.

Definition If $\mathbf{x}$ and $\mathbf{y}$ are nonzero vectors in $\mathbb{R}^n$, then we call

$$\theta = \cos^{-1}\left(\frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|}\right)$$

the *angle between* $\mathbf{x}$ and $\mathbf{y}$.

Example Suppose $\mathbf{x} = (1, 2, 3)$ and $\mathbf{y} = (1, -2, 2)$. Then $\mathbf{x} \cdot \mathbf{y} = 1 - 4 + 6 = 3$, $|\mathbf{x}| = \sqrt{14}$, and $|\mathbf{y}| = 3$, so if θ is the angle between $\mathbf{x}$ and $\mathbf{y}$, we have

$$\cos(\theta) = \frac{3}{3\sqrt{14}} = \frac{1}{\sqrt{14}}.$$

Hence

$$\theta = \cos^{-1}\left(\frac{1}{\sqrt{14}}\right) = 1.3002,$$

where the final value has been rounded to four decimal places.

Example Suppose $\mathbf{x} = (2, -1, 3, 1)$ and $\mathbf{y} = (-2, 3, 1, -4)$. Then $\mathbf{x} \cdot \mathbf{y} = -8$, $|\mathbf{x}| = \sqrt{15}$, and $|\mathbf{y}| = \sqrt{30}$, so if θ is the angle between $\mathbf{x}$ and $\mathbf{y}$, we have

$$\theta = \cos^{-1}\left(\frac{-8}{\sqrt{15}\sqrt{30}}\right) = 1.9575,$$

where the final value has been rounded to four decimal places.

3.3 Direction angles

Let $\mathbf{x}$ be a vector in $\mathbb{R}^n$ and let α_k, be the angle between $\mathbf{x}$ and the kth axis. We call α_1, α_2, ..., α_n the *direction angles* of $\mathbf{x}$. Now α_k is the angle between $\mathbf{x}$ and the standard basis vector $\mathbf{e}_k$. Thus

$$\cos(\alpha_k) = \frac{\mathbf{x} \cdot \mathbf{e}_k}{|\mathbf{x}||\mathbf{e}_k|} = \frac{x_k}{|\mathbf{x}|}.$$

We call $\cos(\alpha_1)$, $\cos(\alpha_2)$, ..., $\cos(\alpha_n)$ the *direction cosines* of $\mathbf{x}$.

Example If $\mathbf{x} = (3, 1, 2)$in $\mathbb{R}^3$, then $|\mathbf{x}| = \sqrt{14}$ and the direction cosines of $\mathbf{x}$ are

$$\cos(\alpha_1) = \frac{3}{\sqrt{14}},$$

$$\cos(\alpha_2) = \frac{1}{\sqrt{14}},$$

and

$$\cos(\alpha_3) = \frac{2}{\sqrt{14}}.$$

Hence, to four decimal places,

$$\alpha_1 = 0.6405,$$

$$\alpha_2 = 1.3002,$$

and

$$\alpha_3 = 1.0069.$$

3.4 Orthogonality and projections

Note that if $\mathbf{x}$ and $\mathbf{y}$ are nonzero vectors in $\mathbb{R}^n$ with $\mathbf{x} \cdot \mathbf{y} = 0$, then the angle between $\mathbf{x}$ and $\mathbf{y}$ is

$$\cos^{-1}(0) = \frac{\pi}{2}.$$

Definition Vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ are said to be *orthogonal* (or *perpendicular*), denoted $\mathbf{x} \perp \mathbf{y}$, if $\mathbf{x} \cdot \mathbf{y} = 0$.

It is a convenient convention of mathematics not to restrict the definition of orthogonality to nonzero vectors. Hence it follows from the definition that $\mathbf{0}$ is orthogonal to every vector in $\mathbb{R}^n$. Moreover, $\mathbf{0}$ is the only vector in $\mathbb{R}^n$ which has this property.

Example The vectors $\mathbf{x} = (-1, -2)$ and $\mathbf{y} = (1, 2)$ are both orthogonal to $\mathbf{z} = (2, -1)$ in $\mathbb{R}^2$. Note that $\mathbf{y} = -\mathbf{x}$ and, in fact, any scalar multiple of $\mathbf{x}$ is orthogonal to $\mathbf{z}$.

Example In $\mathbb{R}^4$, $\mathbf{x} = (1, -1, 1, -1)$ is orthogonal to $\mathbf{y} = (1, 1, 1, 1)$. As in the previous example, any scalar multiple of $\mathbf{x}$ is orthogonal to $\mathbf{y}$.

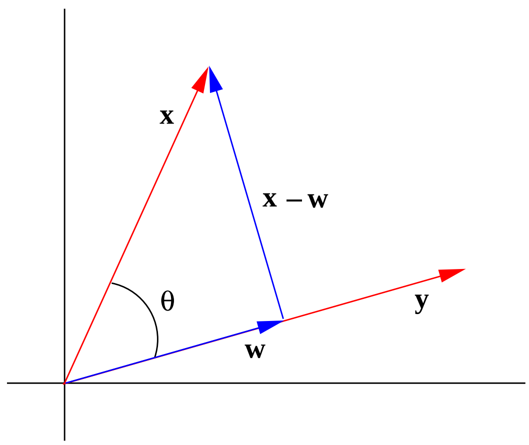
Perhaps the most important application of the dot product is in finding the orthogonal projection of one vector onto another. This is illustrated in the figure below, where $\mathbf{w}$ represents the projection of $\mathbf{x}$ onto $\mathbf{y}$. The result of the projection is to break $\mathbf{x}$ into the sum of two vectors, $\mathbf{w}$, which is parallel to $\mathbf{y}$, and $\mathbf{x} - \mathbf{w}$, which is orthogonal to $\mathbf{y}$, a procedure which is frequently very useful. To compute $\mathbf{w}$, note that if θ is the angle between $\mathbf{x}$ and $\mathbf{y}$, then

$$|\mathbf{w}| = |\mathbf{x}||\cos(\theta)| = |\mathbf{x}|\frac{|\mathbf{x}\cdot\mathbf{y}|}{|\mathbf{x}||\mathbf{y}|} = \left|\mathbf{x}\cdot\frac{\mathbf{y}}{|\mathbf{y}|}\right| = |\mathbf{x}\cdot\mathbf{u}|,$$

where

$$\mathbf{u} = \frac{\mathbf{y}}{|\mathbf{y}|}$$

is the unit vector in the direction of $\mathbf{y}$. Noting that $\mathbf{w}$ has the opposite direction of $\mathbf{y}$ when $\theta > \frac{\pi}{2}$, that is, when $\mathbf{x}\cdot\mathbf{u} < 0$, we see that $\mathbf{w} = (\mathbf{x}\cdot\mathbf{u})\mathbf{u}$.



The projection of $\mathbf{x}$ onto $\mathbf{y}$

Definition Given vectors $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{y} \neq \mathbf{0}$, in $\mathbb{R}^n$, the vector

$$\text{proj}_{\mathbf{y}}\mathbf{x} = (\mathbf{x}\cdot\mathbf{u})\mathbf{u},$$

where $\mathbf{u}$ is the unit vector in the direction of $\mathbf{y}$, is called the *vector projection*, or simply *projection*, of $\mathbf{x}$ onto $\mathbf{y}$. We call

$$\text{comp}_{\mathbf{y}}\mathbf{x} = \mathbf{x}\cdot\mathbf{u}$$

the *scalar projection* of $\mathbf{x}$ onto $\mathbf{y}$.

Example Suppose $\mathbf{x} = (1,2,3)$ and $\mathbf{y} = (1,4,0)$. Then the unit vector in the direction of $\mathbf{y}$ is

$$\mathbf{u} = \frac{1}{\sqrt{17}}(1,4,0),$$

so the scalar projection of $\mathbf{x}$ onto $\mathbf{y}$ is

$$\text{comp}_{\mathbf{y}}\mathbf{x} = \mathbf{x}\cdot\mathbf{u} = \frac{1}{\sqrt{17}}(1+8+0) = \frac{9}{\sqrt{17}}.$$

Thus the vector projection of $\mathbf{x}$ onto $\mathbf{y}$ is

$$\text{proj}_{\mathbf{y}}\mathbf{x} = \frac{9}{\sqrt{17}}\mathbf{u} = \frac{9}{17}(1,4,0) = \left(\frac{9}{17}, \frac{36}{17}, 0\right).$$