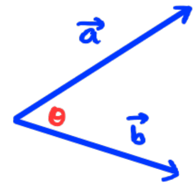
Goal: dot products, angles between vectors, correlation coefficients

Proposition 2.1.1. The **angle** $0° \leq \theta \leq 180°$ between two nonzero 2-vectors $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ satisfies

$$\cos\theta = \frac{a_1 b_1 + a_2 b_2}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}.$$



This is actually the Law of Cosines! (This is worked out on page 24 of the textbook.)

Example 1. Find a nonzero vector in $\mathbb{R}^2$ that is perpendicular to $\mathbf{u} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$.

If $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ were perpendicular to $\vec{u}$, then $\cos\theta = \cos 90° = 0$. Then,

$$0 = \cos 90° = \frac{4v_1 - 3v_2}{\|\vec{u}\| \cdot \|\vec{v}\|} \Rightarrow 4v_1 - 3v_2 = 0.$$

There are infinitely many such vectors, depending on our choice of v_1 (or v_2). For example, picking $v_1 = 3$, we see that $\vec{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ is a vector perpendicular to $\vec{u}$.

Consider two n-vectors $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$.

1. The **dot product** of $\mathbf{x}$ and $\mathbf{y}$ is defined to be the *scalar*

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

2. The angle θ between $\mathbf{x}$ and $\mathbf{y}$, assuming both are nonzero vectors, satisfy the equation

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

Numerator in Prop 2.1.1 is $\vec{a} \cdot \vec{b}$.

3. When $\mathbf{x} \cdot \mathbf{y} = 0$ (for nonzero $\mathbf{x}$ and $\mathbf{y}$), we say that $\mathbf{x}$ and $\mathbf{y}$ are **perpendicular** or **orthogonal**.

Example 2. Find the angle between $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$.

We compute

$$\cos\theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \cdot \|\vec{v}\|} = \frac{1 \cdot 0 + 2 \cdot 3 + (-3) \cdot 1}{\sqrt{14} \cdot \sqrt{10}} = \frac{3}{\sqrt{140}}.$$

Hence, the angle is $\theta = \cos^{-1}\left(\frac{3}{\sqrt{140}}\right) = 75.31°$

Example 3. Show that, for an n-vector $\mathbf{v}$, $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$.

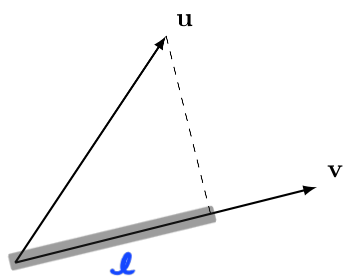
Note that the angle between $\vec{v}$ and $\vec{v}$ is $0°$. Hence,

$$1 = \cos 0° = \frac{\vec{v} \cdot \vec{v}}{\|\vec{v}\| \cdot \|\vec{v}\|} = \frac{\vec{v} \cdot \vec{v}}{\|\vec{v}\|^2},$$

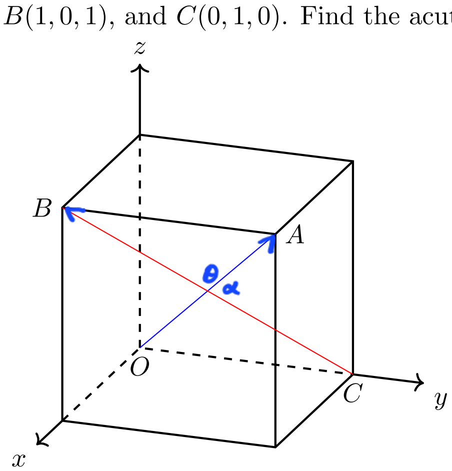
and so, $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$.

We will go over this in more detail in Chapter 6 (Tuesday, June 28), but there is a physical interpretation of the dot product that we can accept for now. Given two non-zero n-vectors $\mathbf{u}$ and $\mathbf{v}$, we can think of the "shadow" that $\mathbf{u}$ projects onto the ray containing $\mathbf{v}$. If the length of the "shadow" is l, then

$$\mathbf{u} \cdot \mathbf{v} = l\|\mathbf{v}\|. \longleftarrow l = \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|}$$



Example 4. Consider the following unit cube in three dimensions and four vertices $O(0,0,0)$, $A(1,1,1)$, $B(1,0,1)$, and $C(0,1,0)$. Find the acute angle between two main diagonals $\overline{OA}$ and $\overline{BC}$.

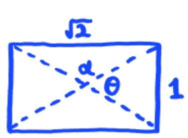


$$\vec{OA} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } \vec{CB} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

$$\cos\theta = \frac{\vec{OA} \cdot \vec{CB}}{\|\vec{OA}\| \|\vec{CB}\|} = \frac{1}{3} \Rightarrow \theta = 70.53°.$$

If you take $\vec{OA}$ and $\vec{BC}$ (instead of $\vec{CB}$), you would get $\alpha = 109.47°$, the supplementary angle.

Food for thought: if you pick a different pair of main diagonals, does the angle change? (the answer is no; why?)



Every pair of diagonals of the unit cube is the pair of diagonals of a $1 \times \sqrt{2}$ rectangle. Thus, the angles do not change.

Properties of dot product (Theorem 2.2.1). Let $\mathbf{v}$, $\mathbf{w}$, $\mathbf{w}_1$, and $\mathbf{w}_2$ be n-vectors and c a scalar. Then,

(a) $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$

(b) $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$

(c) $\mathbf{v} \cdot (c\mathbf{w}) = c(\mathbf{v} \cdot \mathbf{w})$

(d) $\mathbf{v} \cdot (\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{v} \cdot \mathbf{w}_1 + \mathbf{v} \cdot \mathbf{w}_2$

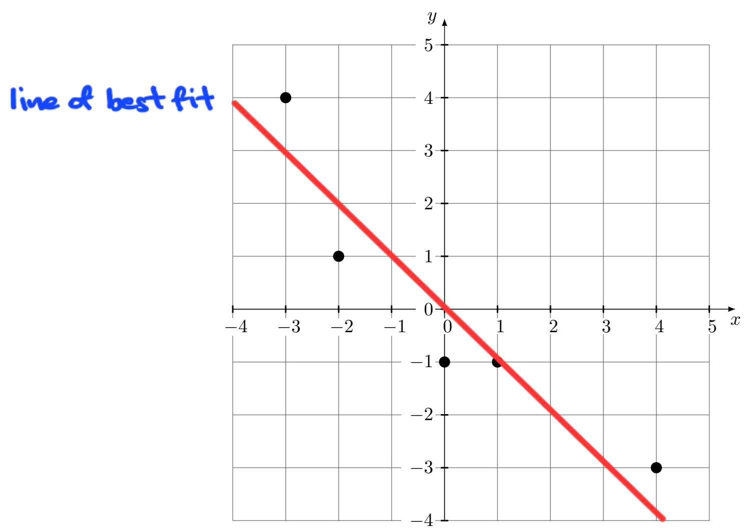
(e) $\mathbf{v} \cdot (c_1\mathbf{w}_1 + c_2\mathbf{w}_2) = c_1(\mathbf{v} \cdot \mathbf{w}_1) + c_2(\mathbf{v} \cdot \mathbf{w}_2)$

Given data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, it is often useful to seek a "best fit line" to this data. But before trying to find the best fit line, it may be more prudent to determine if it is worth finding such a line – i.e. is there a resemblance of a relation between the x_i's and the y_i's? The *correlation coefficient* measures the "strength" of a linear relation between the x_i's and the y_i's.

Let us work with the data set

$$(-3, 4) \qquad (-2, 1) \qquad (0, -1) \qquad (1, -1) \qquad (4, -3)$$

When you plot the data in the xy-plane, they look like:



We can create two vectors – one containing the x_i's and the other containing the y_i's, where the corresponding components come from the same data point:

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ -1 \\ -1 \\ -3 \end{bmatrix} \qquad \bar{x} = \bar{y} = 0$$

We assume that all the data points do not lie on a single horizontal line (i.e. all y_i's are equal) or a single vertical line (i.e. all x_i's are equal). In this running example, the averages of the x_i's and the y_i's are both zero. In the cases they are not (they usually are non-zero), there is a way of shifting them (by subtracting the x-average from the x_i's and the y-average from the y_i's) so that the new averages are both zero.

In this class, for correlation coefficient questions, $\bar{x}$ and $\bar{y}$ will always be 0. For general data sets, define $\hat{\vec{X}} = \vec{X} - \bar{x}\vec{1}$ and $\hat{\vec{Y}} = \vec{Y} - \bar{y}\vec{1}$. Then,

$$r = \frac{\hat{\vec{X}} \cdot \hat{\vec{Y}}}{\|\hat{\vec{X}}\| \|\hat{\vec{Y}}\|}$$

Assuming that the averages of the x_i's and y_i's are both zero, the **correlation coefficient** r between the x_i's and the y_i's are defined as the cosine of the angle between $\mathbf{X}$ and $\mathbf{Y}$, or equivalently, between the unit vectors $\mathbf{X}/\|\mathbf{X}\|$ and $\mathbf{Y}/\|\mathbf{Y}\|$:

$$r := \frac{\mathbf{X}\cdot\mathbf{Y}}{\|\mathbf{X}\|\|\mathbf{Y}\|} = \frac{\mathbf{X}}{\|\mathbf{X}\|}\cdot\frac{\mathbf{Y}}{\|\mathbf{Y}\|}.$$

Example 5. Find the correlation coefficient associated with the data set above.

$$r = \frac{\vec{X}\cdot\vec{Y}}{\|\vec{X}\|\|\vec{Y}\|} = \frac{-27}{\sqrt{30}\sqrt{28}} = \boxed{-0.9316}$$

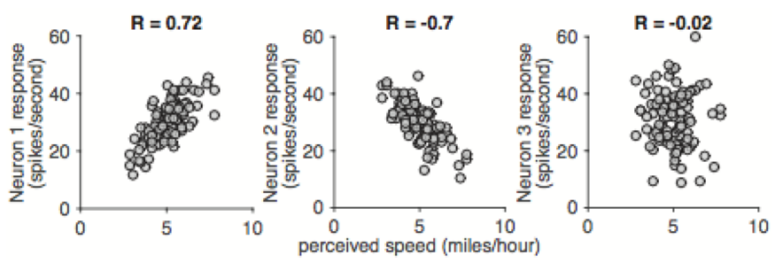
Since magnitude is very close to 1, the line of best fit fits the data very well.

Line of best fit has negative slope.

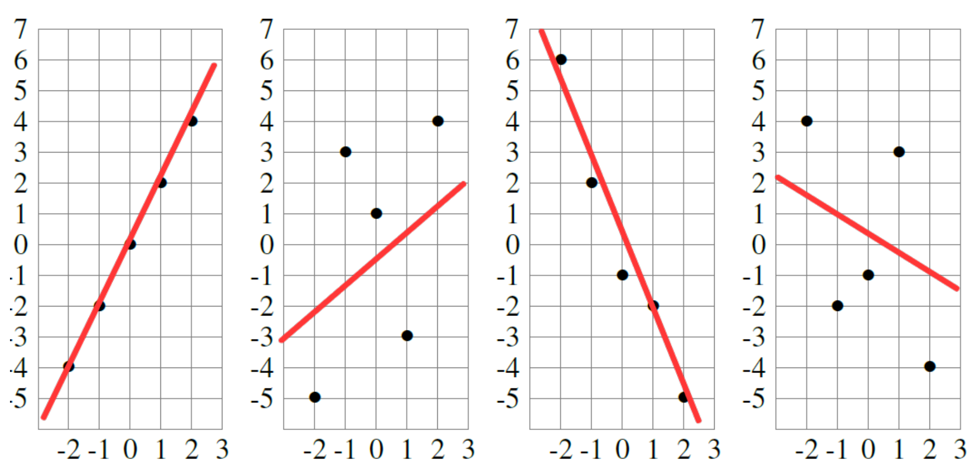
These are evident on the picture on previous page.

Theorem 2.4.5. The correlation coefficient r is always between -1 and 1. When r is close to 1, this means that the data points (x_i, y_i) are close to a line of positive slope, and when r is close to -1, this means that the data points (x_i, y_i) are close to a line of negative slope. A correlation coefficient close to 0 means that there *does not* appear to be a strong linear relation.

Here are some data plots with corresponding correlation coefficients:



Here are some more data plots with corresponding correlation coefficients:



$r = 1$ $\quad$ $r \approx .490$ $\quad$ $r \approx -.983$ $\quad$ $r \approx -.513$

Example 6.

(a) Find all nonzero vectors perpendicular to $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$ and describe geometrically what the collection of all vectors perpendicular to $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$ looks like.

Suppose $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is perpendicular to $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$. Then, $5v_1 + v_2 = 0$, and so,

$$\vec{v} = \begin{bmatrix} v_1 \\ -5v_1 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ -5 \end{bmatrix}.$$

Hence, every scalar multiple of $\begin{bmatrix} 1 \\ -5 \end{bmatrix}$ is perpendicular to $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

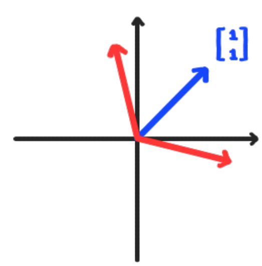
(b) Find all unit vectors that form an angle of 60° with $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

If $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is a unit vector forming an angle of $60°$ with $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$,

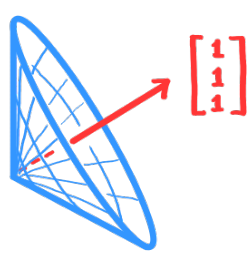
$$\frac{1}{2} = \cos 60° = \frac{v_1 + v_2}{\sqrt{2}} \quad \Rightarrow \quad v_1 + v_2 = \frac{1}{\sqrt{2}}.$$

Plugging this into $v_1^2 + v_2^2 = 1$ gives

$$\left(\frac{\sqrt{2}}{4} - \frac{\sqrt{6}}{4}, \frac{\sqrt{2}}{4} + \frac{\sqrt{6}}{4}\right) \text{ and } \left(\frac{\sqrt{2}}{4} + \frac{\sqrt{6}}{4}, \frac{\sqrt{2}}{4} - \frac{\sqrt{6}}{4}\right).$$



(c) Describe geometrically what the collection of all vectors forming an angle of 60° with $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ looks like.



infinite cone

Example 7. Suppose $\mathbf{v}_1$ and $\mathbf{v}_2$ are unit vectors.

(a) If $\mathbf{v}_1 \perp \mathbf{v}_2$, can we compute $\|4\mathbf{v}_1 + 3\mathbf{v}_2\|$?

Recall that, for any vector $\vec{v}$, $\|\vec{v}\|^2 = \vec{v} \cdot \vec{v}$. Hence,

$$\begin{aligned} \|4\vec{v}_1 + 3\vec{v}_2\|^2 &= (4\vec{v}_1 + 3\vec{v}_2) \cdot (4\vec{v}_1 + 3\vec{v}_2) \\ &= 16\,\vec{v}_1 \cdot \vec{v}_1 + 12\vec{v}_1 \cdot \vec{v}_2 + 12\vec{v}_2 \cdot \vec{v}_1 + 9\,\vec{v}_2 \cdot \vec{v}_2 \\ &= 16\,\|\vec{v}_1\|^2 + 9\,\|\vec{v}_2\|^2 \\ &= 16 + 9 \\ &= 25. \end{aligned}$$

Hence, $\|4\vec{v}_1 + 3\vec{v}_2\| = \boxed{5}.$

(b) If $\mathbf{v}_1 \cdot \mathbf{v}_2 = \dfrac{1}{3}$, can we compute $\|4\mathbf{v}_1 + 3\mathbf{v}_2\|$?

Similarly as in (a),

$$\begin{aligned} \|4\vec{v}_1 + 3\vec{v}_2\|^2 &= 16\|\vec{v}_1\|^2 + 24\,\vec{v}_1 \cdot \vec{v}_2 + 9\|\vec{v}_2\|^2 \\ &= 16 + 24\left(\tfrac{1}{3}\right) + 9 \\ &= 33. \end{aligned}$$

Hence, $\|4\vec{v}_1 + 3\vec{v}_2\| = \boxed{\sqrt{33}}.$

Note that in both (a) and (b), the actual components of $\vec{v}_1$ and $\vec{v}_2$ do not matter.