

Vectors

Last modified: 2026-10-27 20:53:54 (PDT)

Vectors

Exercise 0.1 (Store a data point as one object). A house has a floor area of 120 square meters, 3 bedrooms, and an age of 15 years.

1. Write these three numbers as a single vertical list $\tilde{x}$, in the order given, and call its entries x_1 , x_2 and x_3 . What is x_2 ?
2. Does the list $(3, 120, 15)$ describe the same house?

Solution

Solution 0.1.

1. Stacking the three numbers in order gives

$$\tilde{x} = \begin{bmatrix} 120 \\ 3 \\ 15 \end{bmatrix},$$

so $x_2 = 3$, the number of bedrooms.

2. No. The list $(3, 120, 15)$ would describe a house with 3 square meters of floor area and 120 bedrooms. An entry's position says what it measures, so the order of the entries matters.

Definition 0.1 (Column vector). A **column vector** [1] of length p is an ordered list of p [real numbers](#), written vertically as a $p \times 1$ array:

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

i Remark (Vectors are columns by default)

Remark 0.1 (Vectors are columns by default). In these notes, as in most statistics textbooks, a vector is a column vector unless stated otherwise. A vector written inline as a list in parentheses is still a column vector: for example, $\underset{\sim}{x} = (2, -1, 5)$ means

$$\underset{\sim}{x} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}.$$

i Remark (Why the vector mark sits under the symbol)

Remark 0.2 (Why the vector mark sits under the symbol). In these notes, the mark that shows a symbol is a vector is a tilde under the letter, as in $\underset{\sim}{x}$. Many books instead use a bold letter or an arrow over the letter. We chose a mark underneath because bold letters are hard to write by hand on a board, and because the space above a letter is already used by the hat that marks an estimated value, as in $\hat{\beta}$. The [notation choices page](#) gives the full reasons and an example of an estimated vector.

Definition 0.2 (The set $\mathbb{R}^p$). For a positive [integer](#) p , $\mathbb{R}^p$ is the set of all [column vectors](#) of length p whose entries are [real numbers](#).

i Remark (Reading $\underset{\sim}{x} \in \mathbb{R}^p$)

Remark 0.3 (Reading $\underset{\sim}{x} \in \mathbb{R}^p$). Writing $\underset{\sim}{x} \in \mathbb{R}^p$ says, in one symbol, that $\underset{\sim}{x}$ is a column vector with p real entries. For example, $(1, -2, 0.5) \in \mathbb{R}^3$. The vector $(1, \underset{\sim}{2})$ is not in $\mathbb{R}^3$, because it has only 2 entries; it is in $\mathbb{R}^2$. In [Exercise 0.1](#), $\underset{\sim}{x} \in \mathbb{R}^3$.

Definition 0.3 (Scalar). In linear algebra, a **scalar** [2] is a single [real number](#), as opposed to a vector or a matrix, which holds several numbers.

Example 0.1 (Scalars and a vector). The numbers 5 and -0.2 are scalars. The column vector $(5, -0.2)$ is not a scalar: it has 2 entries, so it is in $\mathbb{R}^2$ ([Definition 0.2](#)), and each of

its 2 entries is a scalar.

Definition 0.4 (Real-valued function, scalar-valued function). A **function** $f : A \rightarrow \mathbb{R}$, whose values are numbers, is **real-valued** (also called **scalar-valued**, since its values are **scalars**).

Definition 0.5 (Vector-valued function). A function $f : A \rightarrow \mathbb{R}^m$, whose values are vectors (Definition 0.2), is **vector-valued**.

Example 0.2 (A real-valued and a vector-valued function). On $\mathbb{R}^2$, $f(\underline{x}) = x_1 + x_2$ is real-valued and $g(\underline{x}) = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \end{bmatrix}$ is vector-valued:

$$f\left(\begin{bmatrix} 3 \\ 4 \end{bmatrix}\right) = 3 + 4 = 7, \quad g\left(\begin{bmatrix} 3 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} 3 + 4 \\ 3 - 4 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}$$

Definition 0.6 (Transpose). The **transpose** of a column vector $\underline{x}$ is the row vector with the same **sequence** of entries, written horizontally:

$$\underline{x}^\top \equiv \underline{x}' \equiv [x_1, x_2, \dots, x_p]$$

Example 0.3 (Transposing a column vector). The transpose of the column vector with entries 2, -1, 5 is the row vector with the same entries:

$$\begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}^\top = [2, -1, 5].$$

Definition 0.7 (Vector addition). The **sum** of two column vectors $\underline{x}$ and $\underline{y}$ of the same length p is the column vector $\underline{x} + \underline{y}$ of length p obtained by adding entry by entry:

$$(\underline{x} + \underline{y})_i \stackrel{\text{def}}{=} x_i + y_i, \quad i = 1, \dots, p$$

Example 0.4 (Adding two vectors).

$$\begin{aligned} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} &= \begin{bmatrix} 1+4 \\ 2+5 \\ 3+6 \end{bmatrix} \\ &= \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix} \end{aligned}$$

Definition 0.8 (Dot product). For any two real-valued vectors $\tilde{x} = (x_1, \dots, x_p)$ and $\tilde{y} = (y_1, \dots, y_p)$ of the same length p , the **dot product** of $\tilde{x}$ and $\tilde{y}$ is:

$$\tilde{x} \cdot \tilde{y} = \tilde{x}^\top \tilde{y} \stackrel{\text{def}}{=} \sum_{i=1}^p x_i y_i \quad (1)$$

See also the definitions in:

- Dobson and Barnett [3], Section 1.3 (equation 1.1, page 7)
- Kaplan [4], chapter on vectors
- [wikipedia](#)

The dot product has a different generalization for two matrices; see [wikipedia](#) for more.

Example 0.5 (A dot product). For $\tilde{x} = (1, 2, 3)$ and $\tilde{y} = (4, 5, 6)$:

$$\begin{aligned} \tilde{x} \cdot \tilde{y} &= 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 && \text{(definition of the dot product)} \\ &= 4 + 10 + 18 && \text{(multiply)} \\ &= 32 && \text{(add)} \end{aligned}$$

Definition 0.9 (Linear combination). A **linear combination** of the numbers $a_1, \dots, a_k$ with coefficients $c_1, \dots, c_k$ is the [sum](#)

$$c_1 a_1 + \dots + c_k a_k = \sum_{i=1}^k c_i a_i$$

A linear combination of vectors $\tilde{v}_1, \dots, \tilde{v}_k$ of the same length is defined entry by entry: entry j of $\sum_{i=1}^k c_i \tilde{v}_i$ is the linear combination of the j th entries of $\tilde{v}_1, \dots, \tilde{v}_k$ with the same coefficients.

Definition 0.10 (Coefficients of a linear combination). The numbers $c_1, \dots, c_k$ in a linear combination (Definition 0.9) $\sum_{i=1}^k c_i a_i$ are its **coefficients**.

Example 0.6 (Coefficients of a linear combination of numbers). The linear combination of the numbers 4 and 5 with coefficients 2 and 3 is $2 \cdot 4 + 3 \cdot 5 = 8 + 15 = 23$.

Definition 0.11 (Weighted sum). A **weighted sum** of the numbers $a_1, \dots, a_k$ is a **linear combination** of them, with **coefficients** $w_1, \dots, w_k$:

$$w_1 a_1 + \dots + w_k a_k = \sum_{i=1}^k w_i a_i$$

Multiply each a_i by its coefficient w_i , then add.

Definition 0.12 (Weight). The **coefficients** $w_1, \dots, w_k$ of a **weighted sum** $\sum_{i=1}^k w_i a_i$ are called its **weights**. The weight w_i says how much a_i counts.

Example 0.7 (The total price of a shopping basket). A basket holds 3 items that cost \$2 each and 2 items that cost \$5 each. The total price is the **weighted sum** of the prices 2 and 5 with **weights** 3 and 2, the count of items at each price:

$$\begin{aligned} 3 \cdot 2 + 2 \cdot 5 &= 6 + 10 && \text{(multiply each price by its weight)} \\ &= 16 && \text{(add)} \end{aligned}$$

Definition 0.13 (Weighted average, weighted mean). A **weighted average** (or **weighted mean**) of the numbers $a_1, \dots, a_k$ is a **weighted sum** of them whose **weights** $w_1, \dots, w_k$ are non-negative and add to 1:

$$\sum_{i=1}^k w_i a_i \quad \text{with} \quad w_i \geq 0 \text{ for each } i \quad \text{and} \quad \sum_{i=1}^k w_i = 1$$

An a_i with a larger **weight** counts for more.

Example 0.8 (A course grade). A course grade counts the midterm for 1/4 and the final for 3/4. A student scores 70 on the midterm and 90 on the final. The **weights** 1/4 and 3/4 are non-negative and add to 1, so the grade is a **weighted average** of 70 and 90:

$$\begin{aligned}\frac{1}{4} \cdot 70 + \frac{3}{4} \cdot 90 &= 17.5 + 67.5 && \text{(multiply each score by its weight)} \\ &= 85 && \text{(add)}\end{aligned}$$

The grade 85 lies between 70 and 90, and nearer to 90, which has the larger [weight](#).

Example 0.9 (A linear combination of two vectors). The linear combination of $(1, 0)^\top$ and $(0, 1)^\top$ with coefficients 2 and 3 is $(2 \cdot 1 + 3 \cdot 0, 2 \cdot 0 + 3 \cdot 1)^\top = (2, 3)^\top$.

i Remark (The dot product is a linear combination)

Remark 0.4 (The dot product is a linear combination). The dot product $\tilde{x} \cdot y$ is a linear combination (Definition 0.9) of the entries of y , with coefficients $x_1, \dots, x_p$. In Example 0.5, $\tilde{x} \cdot y = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6$ is the linear combination of 4, 5, and 6 with coefficients 1, 2, and 3.

The dot product is also the standard *inner product* on $\mathbb{R}^p$; “inner product” is the general notion (Definition 1 in [Inner Products and Orthogonality](#), later in these notes), of which the dot product is one example.

Theorem 0.1 (Dot product is symmetric). *The dot product is symmetric:*

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

i Proof

Proof.

$$\begin{aligned}\tilde{x} \cdot \tilde{y} &= \sum_{i=1}^p x_i y_i && \text{(definition of the dot product)} \\ &= \sum_{i=1}^p y_i x_i && \text{(products of numbers are symmetric)} \\ &= \tilde{y} \cdot \tilde{x} && \text{(definition of the dot product)}\end{aligned}$$

□

Special vectors

Definition 0.14 (Zero vector). The **zero vector** $\mathbf{0}$ of length p has all entries equal to zero:

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

i Remark (The zero vector is the additive identity)

Remark 0.5 (The zero vector is the additive identity). Adding the zero vector to a vector leaves it unchanged (Definition 0.7): $\tilde{x} + \mathbf{0} = \tilde{x}$ for any vector $\tilde{x}$ of the same length, so $\mathbf{0}$ is the **identity element** for vector addition. For example,

$$\begin{aligned} (2, -1) + (0, 0) &= (2 + 0, -1 + 0) \\ &= (2, -1). \end{aligned}$$

Definition 0.15 (Mean, average). The **mean** [5] (or **average**) of n numbers $y_1, \dots, y_n$ is their **sum** divided by how many there are:

$$\bar{y} \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n y_i$$

Example 0.10 (The mean of three numbers). For $y_1 = 2$, $y_2 = -1$ and $y_3 = 5$:

$$\begin{aligned} \bar{y} &= \frac{1}{3} (2 + (-1) + 5) && \text{(definition of the mean, with } n = 3) \\ &= \frac{1}{3} \cdot 6 && \text{(add)} \\ &= 2 && \text{(multiply)} \end{aligned}$$

i Remark (The sample mean estimates the population mean)

Remark 0.6 (The sample mean estimates the population mean). When the numbers are observed data values, $\bar{y}$ is also called the sample mean, an estimator of the **population mean**.

Definition 0.16 (Ones vector). The **ones vector** $\mathbf{1}_{\sim}$ of length p has all entries equal to one:

$$\mathbf{1}_{\sim} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Definition 0.17 (Sum of the elements of a vector). The **sum of the elements** of a vector $x \in \mathbb{R}^p$ is the **sum** of its entries:

$$s_{\sim x} \stackrel{\text{def}}{=} \sum_{i=1}^p x_i.$$

In words, $s_{\sim x}$ is the entries of x added together.

Example 0.11 (The sum of the elements of a vector with three entries). For $x = (2, -1, 5)$,

$$\begin{aligned} s_{\sim x} &= 2 + (-1) + 5 && \text{(definition of the sum of the elements, with } p = 3) \\ &= 6. && \text{(add)} \end{aligned}$$

Theorem 0.2 (The ones vector sums the entries). *For every $x \in \mathbb{R}^p$, the **sum of the elements** of x is the **dot product** of the **ones vector** with x :*

$$s_{\sim x} = \mathbf{1}_{\sim} \cdot x.$$

i Proof

Proof.

$$\begin{aligned} \mathbf{1}_{\sim} \cdot x &= \sum_{i=1}^p 1 \cdot x_i && \text{(Definition 0.8, with every entry of } \mathbf{1}_{\sim} \text{ equal to 1)} \\ &= \sum_{i=1}^p x_i && \text{(multiply each entry by 1)} \\ &= s_{\sim x}. && \text{(Definition 0.17)} \end{aligned}$$

□

Example 0.12 (The ones vector applied to a vector with three entries). For $\tilde{x} = (2, -1, 5)$ and the **ones vector** $\tilde{1}$ with three entries,

$$\begin{aligned}\tilde{1} \cdot \tilde{x} &= 1 \cdot 2 + 1 \cdot (-1) + 1 \cdot 5 \quad (\text{Definition 0.8}) \\ &= 6, \quad (\text{multiply and add})\end{aligned}$$

which is $s_{\tilde{x}}$, as in Example 0.11. Dividing by $p = 3$ gives the **mean** of the entries, $6/3 = 2$.

Definition 0.18 (Indicator vector, standard basis vector). The j -th **indicator vector** (or **standard basis vector** [6]) $\tilde{e}_j$ of length p has a 1 in position j and 0s elsewhere:

$$(\tilde{e}_j)_i = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad \tilde{e}_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{position } j$$

i Remark (Indicator vectors build every vector)

Remark 0.7 (Indicator vectors build every vector). For $p = 3$, the indicator vectors are $\tilde{e}_1 = (1, 0, 0)$, $\tilde{e}_2 = (0, 1, 0)$, and $\tilde{e}_3 = (0, 0, 1)$. Every vector of length 3 is a weighted sum $c_1\tilde{e}_1 + c_2\tilde{e}_2 + c_3\tilde{e}_3$ of them, with weights equal to its entries; for example, $(2, -1, 5) = 2\tilde{e}_1 - 1\tilde{e}_2 + 5\tilde{e}_3$. No other weights work: entry j of $c_1\tilde{e}_1 + c_2\tilde{e}_2 + c_3\tilde{e}_3$ is c_j , so each weight c_j must equal entry j of the vector.

Theorem 0.3 (Indicator vectors select entries). For any vector $\tilde{x}$ of length p and any $j \in \{1, \dots, p\}$:

$$\tilde{e}_j \cdot \tilde{x} = x_j$$

i Proof

Proof. Writing the dot product componentwise:

$$\begin{aligned}
e_{\sim j} \cdot x &= \sum_{i=1}^p (e_{\sim j})_i x_i && \text{(definition of the dot product)} \\
&= \sum_{i=1}^p \begin{cases} 1 \cdot x_i & \text{if } i = j \\ 0 \cdot x_i & \text{if } i \neq j \end{cases} && \text{(definition of } e_{\sim j}) \\
&= x_j && \text{(only the } i = j \text{ term is nonzero)}
\end{aligned}$$

□

Theorem 0.4 (A vector is a linear combination of indicator vectors). *Every vector $x \in \mathbb{R}^p$ is the linear combination (Definition 0.9) of the indicator vectors $e_{\sim 1}, \dots, e_{\sim p}$ (Definition 0.18) with coefficients $x_1, \dots, x_p$:*

$$x_{\sim} = \sum_{j=1}^p x_j e_{\sim j}$$

i Proof

Proof. Compare the two sides entry by entry. For each $i \in \{1, \dots, p\}$:

$$\begin{aligned}
\left(\sum_{j=1}^p x_j e_{\sim j} \right)_i &= \sum_{j=1}^p x_j (e_{\sim j})_i && \text{(Definition 0.9, entry by entry)} \\
&= \sum_{j=1}^p \begin{cases} x_j \cdot 1 & \text{if } j = i \\ x_j \cdot 0 & \text{if } j \neq i \end{cases} && \text{(Definition 0.18)} \\
&= x_i \cdot 1 + \sum_{j \neq i} x_j \cdot 0 && \text{(split off the } j = i \text{ term)} \\
&= x_i + 0 && \text{(times 1 changes nothing; times 0 gives 0)} \\
&= x_i && \text{(adding 0 changes nothing)}
\end{aligned}$$

□

Example 0.13 (Expanding a vector in indicator vectors).

$$\begin{aligned}
\begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix} &= 4 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (-1) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
&= 4e_{\sim 1} - e_{\sim 2} + 2e_{\sim 3}
\end{aligned}$$

Orthogonality

Definition 0.19 (Orthogonal vectors). Two vectors $\tilde{x}$ and $\tilde{y}$ of the same length are **orthogonal** [7] (written $\tilde{x} \perp \tilde{y}$) if their dot product (Definition 0.8) is zero:

$$\tilde{x} \perp \tilde{y} \iff \tilde{x} \cdot \tilde{y} = 0$$

i Remark (Orthogonal means perpendicular)

Remark 0.8 (Orthogonal means perpendicular). Orthogonality extends the geometric idea of perpendicular lines to vectors with any number of entries. In the plane, $(1, 2)$ and $(-2, 1)$ are perpendicular, and their dot product is $1 \cdot (-2) + 2 \cdot 1 = 0$. The same test works in $\mathbb{R}^3$, where pictures are harder to draw:

$$\begin{aligned}(1, 1, 0) \cdot (1, -1, 5) &= 1 - 1 + 0 \\ &= 0,\end{aligned}$$

so $(1, 1, 0) \perp (1, -1, 5)$.

Example 0.14 (Vectors that are not orthogonal).

$$\begin{aligned}(1, 2) \cdot (1, 1) &= 1 \cdot 1 + 2 \cdot 1 \\ &= 3 \\ &\neq 0,\end{aligned}$$

so $(1, 2)$ and $(1, 1)$ are not orthogonal.

Definition 0.20 (Euclidean norm (length, L_2 norm)). The **Euclidean norm** [8] (or **length**, also called the L_2 **norm** and written $\|\tilde{x}\|_2$) of a vector $\tilde{x}$ of length p is

$$\|\tilde{x}\| \stackrel{\text{def}}{=} \sqrt{\tilde{x} \cdot \tilde{x}} = \sqrt{\sum_{i=1}^p x_i^2} \quad (2)$$

where $\sqrt{\cdot}$ is the [square root](#).

Example 0.15 (The length of a vector). For $\tilde{x} = (3, 4)$:

$$\begin{aligned}
\|\tilde{x}\| &= \sqrt{3^2 + 4^2} && \text{(definition of the norm)} \\
&= \sqrt{25} && \text{(square and add)} \\
&= 5
\end{aligned}$$

The vector $(0.6, 0.8)$ has norm $\sqrt{0.36 + 0.64} = 1$.

Definition 0.21 (Unit vector). A **unit vector** [9] is a vector whose Euclidean norm (Definition 0.20) is 1. A unit vector is said to have *unit length*.

Example 0.16 (A unit vector and a vector that is not one). For $\tilde{x} = (0.6, 0.8)$:

$$\begin{aligned}
\|\tilde{x}\| &= \sqrt{0.6^2 + 0.8^2} && \text{(definition of the norm)} \\
&= \sqrt{0.36 + 0.64} && \text{(square)} \\
&= \sqrt{1} && \text{(add)} \\
&= 1 && \text{(take the square root)}
\end{aligned}$$

so $(0.6, 0.8)$ is a unit vector. The vector $(1, 1)$ is not:

$$\begin{aligned}
\|(1, 1)\| &= \sqrt{1^2 + 1^2} \\
&= \sqrt{2} \\
&\neq 1.
\end{aligned}$$

Exercise 0.2 (How far apart are two points?). Let $\tilde{x} = (1, 2)$ and $\tilde{y} = (4, 6)$.

1. Compute the vector $\tilde{x} - \tilde{y}$.
2. Compute the length of $\tilde{x} - \tilde{y}$, using the [Euclidean norm](#).
3. Is the result the same if you compute the length of $\tilde{y} - \tilde{x}$ instead?

Solution

Solution 0.2.

- 1.

$$\begin{aligned}
\tilde{x} - \tilde{y} &= (1 - 4, 2 - 6) \\
&= (-3, -4).
\end{aligned}$$

2.

$$\begin{aligned}\|x - y\| &= \sqrt{(-3)^2 + (-4)^2} \\ &= \sqrt{25} \\ &= 5.\end{aligned}$$

3. $y - x = (3, 4)$, and

$$\begin{aligned}\|y - x\| &= \sqrt{3^2 + 4^2} \\ &= 5.\end{aligned}$$

Yes, the result is the same.

Definition 0.22 (Euclidean distance). The **Euclidean distance** [10] between two vectors x and y of length p is the **Euclidean norm** of their difference:

$$d(x, y) \stackrel{\text{def}}{=} \|x - y\| = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (3)$$

i Remark (Distance from the origin)

Remark 0.9 (Distance from the origin). The distance from the origin 0 to a vector x is the norm of x :

$$\begin{aligned}d(x, 0) &= \|x - 0\| \\ &= \|x\|\end{aligned}$$

[11, p. 39].

For example, the distance from 0 to $(3, 4)$ is $\|(3, 4)\| = 5$, as in Example 0.15.

Definition 0.23 (Orthonormal vectors). A set of vectors $\{x_1, x_2, \dots, x_k\}$ is **orthonormal** if the vectors are mutually orthogonal (Definition 0.19) and each has norm 1 (Definition 0.20), that is, **unit length**:

$$x_i \cdot x_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Example 0.17 (Indicator vectors are orthonormal). The indicator vectors $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_p$ (Definition 0.18) form an orthonormal set. By Theorem 0.3, $\tilde{e}_i \cdot \tilde{e}_j$ is entry i of $\tilde{e}_j$, which is 1 if $i = j$ and 0 if $i \neq j$. For $p = 2$:

$$\begin{aligned}\tilde{e}_1 \cdot \tilde{e}_1 &= 1 \cdot 1 + 0 \cdot 0 \\ &= 1,\end{aligned}$$

$$\begin{aligned}\tilde{e}_2 \cdot \tilde{e}_2 &= 0 \cdot 0 + 1 \cdot 1 \\ &= 1,\end{aligned}$$

and

$$\begin{aligned}\tilde{e}_1 \cdot \tilde{e}_2 &= 1 \cdot 0 + 0 \cdot 1 \\ &= 0.\end{aligned}$$

Example 0.18 (Two ways to fail to be orthonormal).

- $(1, 1)$ and $(1, -1)$ are orthogonal, since $1 \cdot 1 + 1 \cdot (-1) = 0$, but not orthonormal: $(1, 1) \cdot (1, 1) = 2 \neq 1$.
- $(1, 0)$ and $(0.6, 0.8)$ both have unit length, since $(1, 0) \cdot (1, 0) = 1$ and

$$\begin{aligned}(0.6, 0.8) \cdot (0.6, 0.8) &= 0.36 + 0.64 \\ &= 1,\end{aligned}$$

but they are not orthonormal: $(1, 0) \cdot (0.6, 0.8) = 0.6 \neq 0$.

Exercise 0.3 (Compute an inner product and a norm). Let

$$\tilde{a} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} 0 \\ 4 \\ -2 \end{bmatrix}$$

Compute $\tilde{a}^\top \tilde{b}$ and $\|\tilde{a}\|_2$.

Solution

Solution 0.3. Multiply entry by entry and add:

$$\begin{aligned}\tilde{a} \cdot \tilde{b} &= (3)(0) + (-1)(4) + (2)(-2) \\ &= 0 - 4 - 4 \\ &= -8\end{aligned}$$

For the norm, take the inner product of $\tilde{a}$ with itself first:

$$\begin{aligned}\tilde{a} \cdot \tilde{a} &= 3^2 + (-1)^2 + 2^2 \\ &= 9 + 1 + 4 \\ &= 14\end{aligned}$$

so

$$\|\tilde{a}\|_2 = \sqrt{14} \approx 3.742$$

The inner product came out negative while the norm cannot: a norm is a square root of a sum of squares, so it is never below zero.

Python has both operations built in, and R builds them from `sum`.

```
import numpy as np

a = np.array([3.0, -1.0, 2.0])
b = np.array([0.0, 4.0, -2.0])

inner = float(a @ b)
norm_a = round(float(np.linalg.norm(a)), 3)

print(inner, norm_a)
#> -8.0 3.742
```

```
a <- c(3, -1, 2)
b <- c(0, 4, -2)

print(c(sum(a * b), round(sqrt(sum(a^2)), 3)))
#> [1] -8.000 3.742
```

Both print -8 and 3.742.

Video lecture

Hutchinson's [Linear Function Basics](#) (24 min) covers dot products and the Euclidean norm, and goes on to hyperplanes ([Definition 22 in Matrices](#)) and [level sets](#) [12]. The login for the video site is posted [on Canvas](#).

- [1] Wikipedia contributors, “Row and column vectors — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Row_and_column_vectors
- [2] Wikipedia contributors, “Scalar (mathematics) — Wikipedia, the free encyclopedia.” 2026. Available: [https://en.wikipedia.org/wiki/Scalar_\(mathematics\)](https://en.wikipedia.org/wiki/Scalar_(mathematics))
- [3] A. J. Dobson and A. G. Barnett, *An introduction to generalized linear models*, 4th ed. CRC press, 2018. doi: [10.1201/9781315182780](https://doi.org/10.1201/9781315182780).
- [4] D. Kaplan, *MOSAIC calculus*. www.mosaic-web.org, 2022. Available: www.mosaic-web.org
- [5] Wikipedia contributors, “Arithmetic mean — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Arithmetic_mean
- [6] Wikipedia contributors, “Standard basis — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Standard_basis
- [7] Wikipedia contributors, “Orthogonality — Wikipedia, the free encyclopedia.” 2026. Available: <https://en.wikipedia.org/wiki/Orthogonality>
- [8] Wikipedia contributors, “Euclidean vector — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Euclidean_vector
- [9] Wikipedia contributors, “Unit vector — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Unit_vector
- [10] Wikipedia contributors, “Euclidean distance — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Euclidean_distance
- [11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT Press, 2016. Available: <https://www.deeplearningbook.org/>
- [12] B. Hutchinson, “DATA 471/571 (machine learning) and CSCI 481/581 (deep learning) video lectures.” Western Washington University. Accessed: Sep. 28, 2026. [Online]. Available: https://facultyweb.cs.wvu.edu/~hutchib2/video_lectures/data371/