# Matrix-vector multiplication

A vector in $\mathbb{R}^n$ is an ordered $n$-tuple of real numbers. We can write it as a column vector:

$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ or as a row vector $(a_1, a_2, \dots, a_n)$.

Just like with matrices, we can take the sum of two $n$-vectors, and we can take the scalar multiple of a vector.

If $A$ is an $m \times n$ matrix, we can write

$$A = [\vec{a}_1 \ \vec{a}_2 \ \dots \ \vec{a}_n]$$

where each $\vec{a}_i$ is one of its columns. If $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, a vector in $\mathbb{R}^n$, we define the product

$$A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n.$$

Note that we can only multiple a matrix and a vector in $\mathbb{R}^n$ if the matrix has $n$ columns.

**Ex:** $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$.

Then $A\vec{x} = 1\begin{bmatrix}2\\0\end{bmatrix} - 1\begin{bmatrix}3\\1\end{bmatrix} + 3\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}2\\5\end{bmatrix}$.

Notice that we started w/ a vector in $\mathbb{R}^3$ and ended with a vector in $\mathbb{R}^2$.

**Ex:** Let $I = \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$ and $\vec{a} = \begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}$ an arbitrary vector. Then $I\vec{a} = a_1\begin{bmatrix}1\\0\\0\end{bmatrix} + a_2\begin{bmatrix}0\\1\\0\end{bmatrix} + a_3\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix} = \vec{a}$.

So multiplication of $I$ with a vector gives the same vector back. $I$ is called the $3\times 3$ <u>identity matrix</u>.

## Properties of matrix-vector multiplication:

① $\underset{\text{matrix}}{A}(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y}$   vectors

② $A(\underset{\text{scalar}}{a}\vec{x}) = a(A\vec{x}) = (aA)\vec{x}$

③ $(A+B)\vec{x} = A\vec{x} + B\vec{x}$

How does this relate to systems of equations?

Consider the system
$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_1 \qquad - x_3 &= -1 \end{aligned}$$

We can rewrite this as an equality of vectors:
$$\begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ 2x_1 - x_3 \end{bmatrix} = \begin{bmatrix}1\\-1\end{bmatrix}$$, which becomes

$$\begin{bmatrix} x_1 \\ 2x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 3x_3 \\ -x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\Rightarrow \quad x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\Rightarrow \quad \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & -1 \end{bmatrix}}_{\text{coefficient matrix}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{\text{constant matrix (or vector)}}$$

More generally, if we have a system of equations in $n$ variables, $x_1, \dots, x_n$, w/ coefficient matrix $A$, constant vector $\vec{b}$ and set

$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. Then we can express the same system

as $A\vec{x} = \vec{b}$, called the <u>matrix form</u> of the system.

Note that the corresponding augmented matrix is $[A \mid \vec{b}]$.

Suppose $\vec{x}_1$ is a solution to $A\vec{x} = \vec{b}$ and $\vec{x}_0$ is a solution to $A\vec{x} = \vec{0}$, the <u>associated homogeneous system</u>.

$\uparrow$ the $\vec{0}$ vector (each entry is 0)

Then notice the following:

$$A(\vec{x}_1 + \vec{x}_0) = A\vec{x}_1 + A\vec{x}_0 = \vec{b} + \vec{0} = \vec{b}.$$

Thus, $\vec{x}_1 + \vec{x}_0$ is also a solution to $A\vec{x} = \vec{b}$.

In fact, every solution has this form:

**Theorem**: If $\vec{x}_1$ is a solution to the system $A\vec{x} = \vec{b}$, then every solution $\vec{x}_2$ to the system is of the form

$$\vec{x}_2 = \vec{x}_1 + \vec{x}_0$$

where $\vec{x}_0$ is some solution of the associated homogeneous system $A\vec{x} = \vec{0}$.

**Ex**: 
$$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 1 \\ 2x_1 + 3x_2 - x_3 - x_4 &= -2 \end{aligned}$$

$$\left[\begin{array}{cccc|c} 1 & 2 & -3 & 1 & 1 \\ 2 & 3 & -1 & -1 & -2 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & -3 & 1 & 1 \\ 0 & -1 & 5 & -3 & -4 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & -3 & 1 & 1 \\ 0 & 1 & -5 & 3 & 4 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 7 & -5 & -7 \\ 0 & 1 & -5 & 3 & 4 \end{array}\right]$$

$x_1 = -7s + 5t - 7$
$x_2 = 5s - 3t + 4$
$x_3 = s$
$x_4 = t$

General solution: $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -7s + 5t - 7 \\ 5s - 3t + 4 \\ s \\ t \end{bmatrix}$

$$= \underbrace{s\begin{bmatrix} -7 \\ 5 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 5 \\ -3 \\ 0 \\ 1 \end{bmatrix}}_{\vec{x}_0} + \underbrace{\begin{bmatrix} -7 \\ 4 \\ 0 \\ 0 \end{bmatrix}}_{\vec{x}_1}$$

General solution to homogeneous system:

$$s\begin{bmatrix} -7 \\ 5 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 5 \\ -3 \\ 0 \\ 1 \end{bmatrix}$$

(Check this by setting the constants in the original system equal to 0!)

## The dot product

**Def:** If $\vec{a} = (a_1, \dots, a_n)$ and $\vec{b} = (b_1, \dots, b_n)$, two vectors in $\mathbb{R}^n$, the dot product of $\vec{a}$ and $\vec{b}$ is

$$\vec{a} \cdot \vec{b} = a_1b_1 + a_2b_2 + \dots + a_nb_n.$$

The dot product gives us another way to describe matrix-vector multiplication:

e.g. if $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$ and $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ a vector in $\mathbb{R}^4$, Then

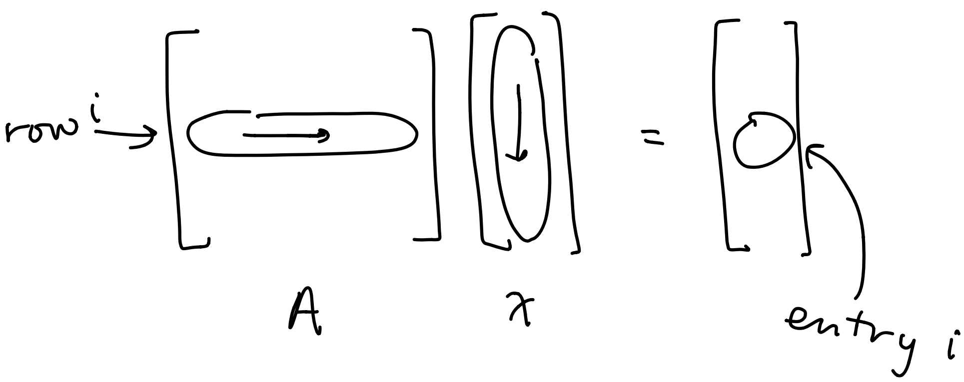
$$A\vec{x} = x_1\begin{bmatrix}1\\5\\9\end{bmatrix} + x_2\begin{bmatrix}2\\6\\10\end{bmatrix} + x_3\begin{bmatrix}3\\7\\11\end{bmatrix} + x_4\begin{bmatrix}4\\8\\12\end{bmatrix}$$

$$= \begin{bmatrix} x_1 + 2x_2 + 3x_3 + 4x_4 \\ 5x_1 + 6x_2 + 7x_3 + 8x_4 \\ 9x_1 + 10x_2 + 11x_3 + 12x_4 \end{bmatrix}$$

each entry is the dot product of $\vec{x}$ with the corresponding row of the matrix

This is true more generally:

**Theorem:** If $A$ is an $m \times n$ matrix and $\vec{x}$ an $n$-vector, then each entry of $A\vec{x}$ is the dot product of the corresponding row of $A$ with $\vec{x}$.




It turns out that matrices are determined by how they multiply with vectors. That is:

**Theorem:** If $A$ and $B$ are $m \times n$ matrices such that for every $n$-vector $\vec{x}$, $A\vec{x} = B\vec{x}$, then $A = B$.

Why? Let $\vec{e}_1 = \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}$, $\vec{e}_2 = \begin{bmatrix}0\\1\\0\\\vdots\end{bmatrix}$, etc. so that $\vec{e}_i$ is the vector with $i^{th}$ entry 1 and remaining entries 0.

Then if $A = [\vec{a}_1 \; \dots \; \vec{a}_m]$, $B = [\vec{b}_1 \; \dots \; \vec{b}_m]$

columns

we have $A\vec{e}_i = B\vec{e}_i$

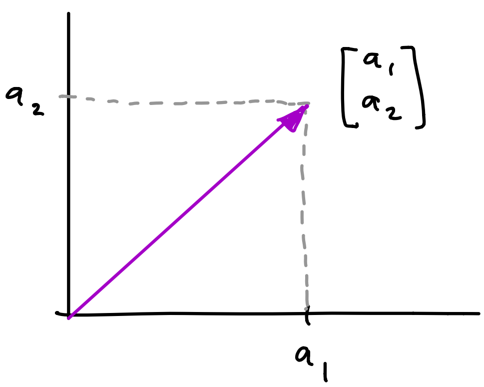
$$\Rightarrow 0\vec{a}_1 \dots + 1\vec{a}_i + \dots + 0a_m = 0\vec{b}_1 + \dots + 1b_i + \dots + 0\vec{b}_m$$

$$\Rightarrow \vec{a}_i = \vec{b}_i.$$

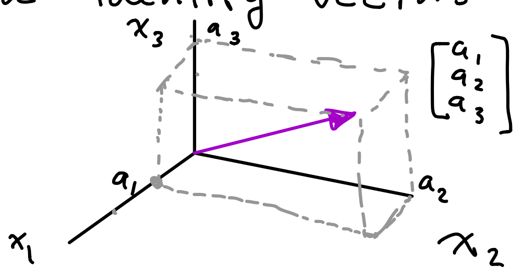
So the $i^{th}$ columns of $A$ and $B$ are the same. But this holds for each $i$, so $A = B$.

## Transformations

One way to think about vectors in $\mathbb{R}^2$ is geometrically as points in the plane. We usually draw them as an arrow from the origin.




Similarly in $\mathbb{R}^3$, we can identify vectors with points in 3-dimensional space.

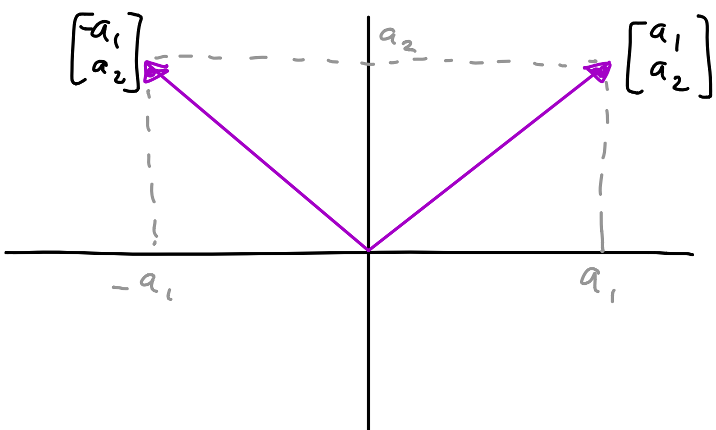


A transformation is a special kind of function from $\mathbb{R}^n$ to $\mathbb{R}^m$ that can be described by a matrix.

Recall that, in general, a function $f$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, written $f: \mathbb{R}^n \to \mathbb{R}^m$, is a rule that assigns each vector in $\mathbb{R}^n$ to a unique vector in $\mathbb{R}^m$.

A function $f: \mathbb{R}^n \to \mathbb{R}^m$ is a transformation if that "rule" is given by multiplying the vector in $\mathbb{R}^n$ by an $m \times n$ matrix to obtain the output in $\mathbb{R}^m$.

**Ex:** Consider the function $T: \mathbb{R}^2 \to \mathbb{R}^2$ that sends every vector to its reflection over the $y$ axis.




So $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ is sent to $\begin{bmatrix} -a_1 \\ a_2 \end{bmatrix}$. That is, $T\left(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}\right) = \begin{bmatrix} -a_1 \\ a_2 \end{bmatrix}$.

But notice that
$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -a_1 \\ a_2 \end{bmatrix}.$$

So for any $\vec{x}$ in $\mathbb{R}^2$, $T(\vec{x}) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \vec{x}$.

This is an example of a transformation. More generally, we have the following definition:

**Def:** A function $T: \mathbb{R}^n \to \mathbb{R}^m$ is a <u>linear transformation</u> if there is some $m \times n$ matrix $A$ such that

$$T(\vec{x}) = A\vec{x}$$

for all $\vec{x}$ in $\mathbb{R}^n$.

Sometimes we can give a formula to describe the transformation as well.

**Ex:** Let $T: \mathbb{R}^3 \to \mathbb{R}^2$ be the function given by

$$T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ x_3 - x_2 \end{bmatrix}.$$

Is this a linear transformation? Yes! Notice that

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ x_3 - x_2 \end{bmatrix}$$

So $T(\vec{x}) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix}\vec{x}$.

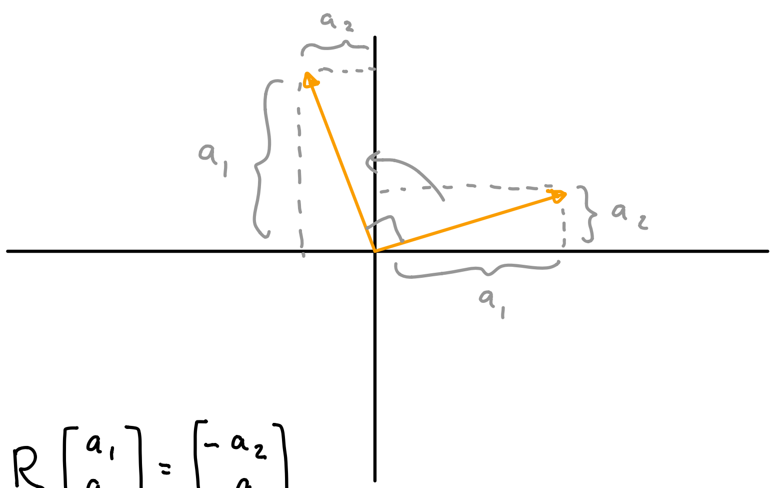
(Soon we'll talk more about finding the matrix that works.)

We can also start with a matrix and use it to define a transformation:

**Def:** If $A$ is an $m \times n$ matrix, the <u>transformation induced by $A$</u>, written $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is defined

$$T_A(\vec{x}) = A\vec{x}, \quad \text{for all } \vec{x} \text{ in } \mathbb{R}^n.$$

**Ex:** Let $R : \mathbb{R}^2 \to \mathbb{R}^2$ denote the function that rotates a vector $90°$ counterclockwise about the origin. Is this a transformation?




Then $R\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -a_2 \\ a_1 \end{bmatrix}$.

We can describe this with a matrix:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -a_2 \\ a_1 \end{bmatrix}$$

So $R$ is the transformation induced by $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

**Def:** 1. The <u>zero transformation</u> $T : \mathbb{R}^n \to \mathbb{R}^m$ is the transformation induced by the $0$ matrix, and is given by

$$T(\vec{x}) = \vec{0}. \quad \text{It is denoted } T = 0.$$

② The identity transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is denoted $T = 1_{\mathbb{R}^n}$, and is given by $T(\vec{x}) = \vec{x}$.

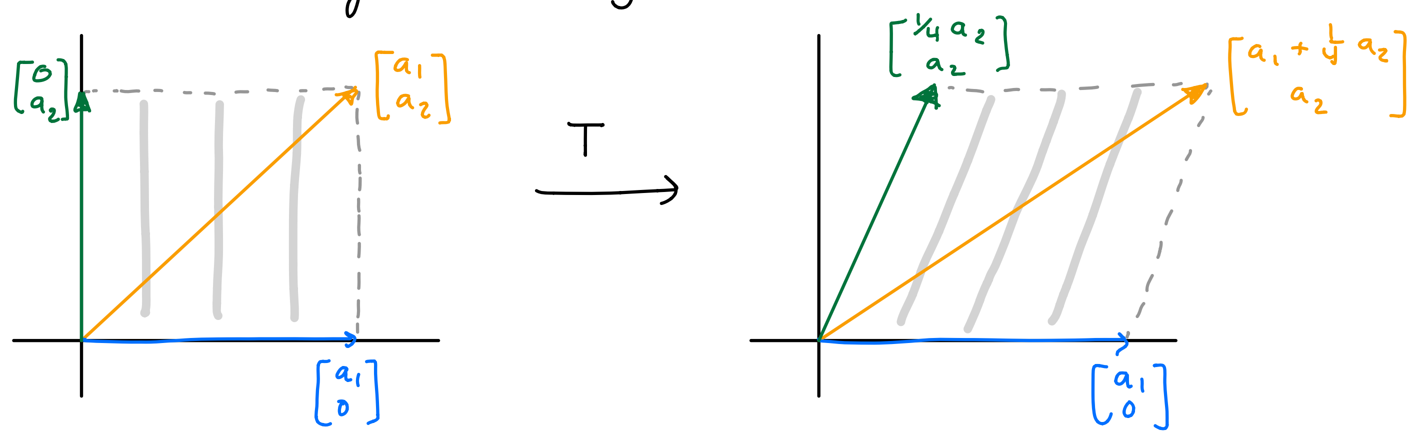
It is the transformation induced by the $n \times n$ identity matrix $I_n = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & \ddots & 1 \end{bmatrix}$ (ones on diagonal, zeros everywhere else).

Ex: If $a$ is a real number, the transformation

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + ay \\ y \end{bmatrix}$$

induced by the matrix $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$ is called an

x-shear of $\mathbb{R}^2$. If, for example $a = \frac{1}{4}$, we can visualize it geometrically:




Note that not every geometric function is a transformation!

Ex: Consider the function $f: \mathbb{R}^2 \to \mathbb{R}^2$ that translates a point one unit to the right. i.e.

$$f\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 1 \\ y \end{bmatrix}.$$

Why isn't this a linear transformation?

Suppose there <u>is</u> some matrix $A$ such that $f(\vec{x}) = A\vec{x}$.

Then, in particular, $f(\vec{0}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = A \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

But $A \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ for any matrix! So this is impossible.

<u>Practice problems</u>: 2.2 : 2, 4, 5ab, 8a, 9, 11bd, 12, 18