4. Matrices and determinants.

4.1. Matrix operations.

We have seen that a <u>matrix</u> is a rectangular array of numbers. An $m \times n$ matrix has m rows and n columns. Two matrices are equal when they have equal size and entries.

<u>Example:</u> A 2×3 matrix is:

$$\begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix},$$

a 3×2 matrix is:

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix},$$

a 2×2 matrix is:

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix},$$

and a general $m \times n$ matrix has the form:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$ where a_{ij} is the entry in row i and column j.

A matrix can be denoted $A = [a_{ij}]$. A <u>column vector</u> is a $m \times 1$ matrix. A <u>row vector</u> is a $1 \times n$ matrix. Let A and B be matrices of size $m \times n$, and let s be a scalar. The matrix $A+B$ is the point-wise addition of the entries of A and B. The matrix sA is the point-wise multiplication of the entries of A by s.

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix}.$$

$$s \cdot \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} sa_{11} & \cdots & sa_{1n} \\ \vdots & & \vdots \\ sa_{m1} & \cdots & sa_{mn} \end{bmatrix}.$$

<u>Example</u>: Let $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$. Then:

$$A + B = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix},$$ but $A+C$ and $B+C$

do <u>not</u> exist because A and C do <u>not</u> have the same shape, and B and C do <u>not</u> have the same shape. Also:

$$2 \cdot \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 10 \\ 4 & 8 & 12 \end{bmatrix} \quad \text{and} \quad 3 \cdot \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 12 \\ 6 & 15 \\ 9 & 18 \end{bmatrix}.$$

Let A be a matrix of size $m \times p$ and B a matrix of size $q \times n$. The __matrix multiplication__ $C = AB$ is defined only when $p = q$, in which case C is the $m \times n$ matrix with entries:

$$c_{ij} = \underbrace{\begin{bmatrix} a_{i1} & \cdots & a_{ip} \end{bmatrix}}_{\text{row } i \text{ of } A} \cdot \underbrace{\begin{bmatrix} b_{1j} \\ \vdots \\ b_{pj} \end{bmatrix}}_{\text{column } j \text{ of } B} = a_{i1} b_{1j} + \cdots + a_{ip} b_{pj} = \sum_{k=1}^{p} a_{ik} b_{kj} .$$

Schematically:

$$\begin{array}{cccc} A & B & = & C \\ m \times p & p \times n & & m \times n \end{array}$$

$$\begin{bmatrix} a_{11} & \cdots & & \cdots & a_{1p} \\ \vdots & & & & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{ip} \\ \vdots & & & & \vdots \\ a_{m1} & \cdots & & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ & & \vdots & & \\ \vdots & & b_{kj} & & \vdots \\ & & \vdots & & \\ b_{p1} & \cdots & b_{pj} & \cdots & b_{pn} \end{bmatrix} = \begin{bmatrix} & & \\ & \boxed{c_{ij}} & \\ & & \end{bmatrix} .$$

__Example:__ Let $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$. Then:

$$AC = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 1+6+15 & 4+15+30 \\ 2+8+18 & 8+20+36 \end{bmatrix} = \begin{bmatrix} 22 & 49 \\ 28 & 64 \end{bmatrix},$$

$$CA = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 1+8 & 3+16 & 5+24 \\ 2+10 & 6+20 & 10+30 \\ 3+12 & 9+24 & 15+36 \end{bmatrix} = \begin{bmatrix} 9 & 19 & 29 \\ 12 & 26 & 40 \\ 15 & 33 & 51 \end{bmatrix},$$

and both AB and BA are <u>not</u> defined because the number of columns of the first matrix is different to the number of rows of the second matrix.

In particular, matrix multiplication is <u>not</u> commutative. We may also have non-zero matrices that multiplied give the zero matrix.

<u>Example</u>: Find a, b, c, d such that: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Writing out the multiplication gives:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2+bc & ab+bd \\ ca+dc & cb+d^2 \end{bmatrix}$$

which gives the following four equations, one for each entry.

$$\begin{cases} a^2+bc=0 \\ b(a+d)=0 \\ c(a+d)=0 \\ d^2+bc=0 \end{cases}$$

Thus $a=b=c=d=0$ is a solution. If we want $a\neq 0$ then we must have $a^2+bc=0$, so $bc=-a^2$, so $b\neq 0\neq c$ and $b=\frac{-a^2}{c}$, and we must have $b(a+d)=0$, so $a+d=0$, so $d=-a$. Thus a is a free non-zero parameter, c is a free non-zero parameter, $b=-\frac{a^2}{c}$, and $d=-a$. If we want $a=0$

then we must have $a^2+bc=0$, so $bc=0$, so either $b=0$ or $c=0$, say $b=0$ and $c\neq 0$, then we must have $c(a+d)=0$ so $d=0$. Thus $a=b=d=0$ and c is a free non-zero parameter. Using that the equations are symmetric when exchanging the roles of a and d, as well as the roles of b and c, we find two more solutions: when $a=c=d=0$ and b is a free non-zero parameter, and when a is a free non-zero parameter, b is a free non-zero parameter, $c=\frac{-a^2}{b}$, and $d=-a$. All solutions have the following form:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix}, \begin{bmatrix} a & b \\ \frac{-a^2}{b} & -a \end{bmatrix}, \begin{bmatrix} a & \frac{-a^2}{c} \\ c & -a \end{bmatrix},$$

but we should note that the last two forms are equivalent (since $c=\frac{-a^2}{b}$ if and only if $b=\frac{-a^2}{c}$).

An equality of matrices is a way of expressing many equalities simultaneously. This will be very useful to condense the information within a system of equations, involving many equations that need to be satisfied at the same time, into one single equality involving matrices. A system of equations corresponds to an equality of matrices:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

corresponds to:

$$\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Denoting:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \ \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \text{ and } \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

a system of equations is just an equality of the form $A\vec{x} = \vec{b}$.

Example: The system of equations:

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 2 \\ 4x_1 + 5x_2 + 6x_3 = 3 \end{cases}$$ corresponds to the equality $$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

More concretely, multiplying a matrix by a column vector, in that order from left to right, gives a linear combination of the column of the matrix. The scalars are the entries of the column vector.

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix} = \begin{bmatrix} a_{11}s_1 + \cdots + a_{1n}s_n \\ \vdots \\ a_{m1}s_1 + \cdots + a_{mn}s_n \end{bmatrix} = \begin{bmatrix} a_{11}s_1 \\ \vdots \\ a_{m1}s_1 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}s_n \\ \vdots \\ a_{mn}s_n \end{bmatrix} =$$

$$= s_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + s_n \begin{bmatrix} a_{1n} \\ \\ a_{mn} \end{bmatrix}.$$

<u>Example</u>: The multiplication:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \\ 32 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} + 2 \cdot \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} + 3 \cdot \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$$

is indeed a linear combination of the columns of the matrix.

4.2. Linear transformations and matrices

A <u>function</u> $f : X \longrightarrow Y$ from the set X to the set Y is an assignment of exactly one element $f(x)$ of Y for each element x in X. The set X is called the <u>domain</u> or <u>source</u> of the function. The set Y is called the <u>codomain</u> or <u>target</u> of the function. Given x in X, the element $f(x)$ in Y is called the <u>image</u> of x. The collection of all elements y in Y such that there is an x in X with $y = f(x)$ is called the <u>image of the function</u> or the <u>range of the function</u>. Plainly, the source or domain is the collection of possible inputs, the target or codomain is the collection of possible outputs, and the image or range is the collection of actual outputs.

Example: The following assignments are functions.

1) $f : \mathbb{R} \longrightarrow \mathbb{R}$ for m a real value.
$x \longmapsto mx$

Namely $f(x) = mx$ where x is any real value.

2) $f : \mathbb{R} \longrightarrow \mathbb{R}$ for m and b real values.
$x \longmapsto mx + b$

Namely $f(x) = mx + b$ where x is any real value.

3) $f : \mathbb{R} \longrightarrow \mathbb{R}$. Namely $f(x) = x^2$ where x is any real value.
$x \longmapsto x^2$

4) $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. Namely $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$.
$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \longmapsto \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$

We will be interested in functions that are linear, namely $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ satisfying:

(i) $f(r \cdot \vec{x}) = r \cdot f(\vec{x})$ for all scalars r and all vectors $\vec{x}$, and

(ii) $f(\vec{x}_1 + \vec{x}_2) = f(\vec{x}_1) + f(\vec{x}_2)$ for all vectors $\vec{x}_1$ and $\vec{x}_2$.

Equivalently, a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear when $f(r \cdot \vec{x}_1 + s \cdot \vec{x}_2) = r \cdot f(\vec{x}_1) + s \cdot f(\vec{x}_2)$.

A linear function is also called a linear transformation.

Example: Determine which of the following functions are linear.

1) $f: \mathbb{R} \longrightarrow \mathbb{R}$, $x \longmapsto mx$ for m a real value is linear.

2) $f: \mathbb{R} \longrightarrow \mathbb{R}$, $x \longmapsto mx+b$ for m and b real values is linear if $b=0$, and it is not linear if $b \neq 0$.

3) $f: \mathbb{R} \longrightarrow \mathbb{R}$, $x \longmapsto x^2$ is not linear.

4) $f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is linear.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \longmapsto \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

In general, a function $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear if and only if there exists an $m \times n$ matrix A such that $f(\vec{x}) = A\vec{x}$.

Given an $m \times n$ matrix, the function $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $\vec{x} \longmapsto A\vec{x}$ is linear, as follows.

(i) $f(r \cdot \vec{x}) = A(r \cdot \vec{x}) = (A \cdot r)\vec{x} = (r \cdot A)\vec{x} = r \cdot (A\vec{x}) = r \cdot f(\vec{x})$

commutativity of scalar multiplication

associativity of matrix multiplication and multiplication by a scalar

(ii) $f(\vec{x}_1 + \vec{x}_2) = A(\vec{x}_1 + \vec{x}_2) = A\vec{x}_1 + A\vec{x}_2 = f(\vec{x}_1) + f(\vec{x}_2)$

distributivity of matrix multiplication and matrix addition

Given a function $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$, it is fully determined by its values on the standard basis vectors $\vec{e}_1, \dots, \vec{e}_n$ because the image $f(\vec{x})$ of any vector $\vec{x}$ can be written as follows.

$$f(\vec{x}) = f\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = f\left(\begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_n \end{bmatrix}\right) = f(x_1 \vec{e}_1 + \dots + x_n \vec{e}_n) \underset{\uparrow \text{ f is linear}}{=} x_1 f(\vec{e}_1) + \dots + x_n f(\vec{e}_n) =$$

$$= \underbrace{\begin{bmatrix} f(\vec{e}_1) & \cdots & f(\vec{e}_n) \end{bmatrix}}_{m\times n \text{ matrix with rows } f(\vec{e}_1), \dots, f(\vec{e}_n)} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

In particular, setting A to be the $m\times n$ matrix with columns $f(\vec{e}_1), \dots, f(\vec{e}_n)$, we have $f(\vec{x}) = A\vec{x}$ for all vectors $\vec{x}$ in $\mathbb{R}^n$.

<u>Example:</u> Find a matrix A such that the linear transformation $f: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \longmapsto \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{bmatrix}$ can be expressed as $f(\vec{x}) = A\vec{x}$.

As we saw above, the matrix associated to a linear transformation has columns the vectors $f(\vec{e}_1), f(\vec{e}_2), f(\vec{e}_3)$. These are:

$$f(\vec{e}_1) = f\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1+0+0 \\ 1+0+0 \\ 4+0+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$$

$$f(\vec{e}_2) = f\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0+1+0 \\ 0+2+0 \\ 0+5+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

$$f(\vec{e}_3) = f\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0+0+1 \\ 0+0+3 \\ 0+0+6 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$$

so the desired matrix is:

$$A = \begin{bmatrix} f(\vec{e}_1) & f(\vec{e}_2) & f(\vec{e}_3) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}.$$

We can check that indeed:

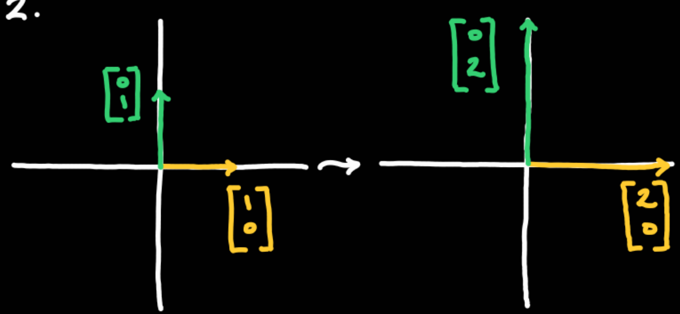
$$A\vec{x} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{bmatrix} = f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = f(\vec{x}).$$

Remark: An important conceptual significance of the meaning behind systems of linear equations is: given a matrix A and a vector $\vec{b}$, asking for solutions $\vec{x}$ to the system of linear equations $A\vec{x} = \vec{b}$ is equivalent to asking if the point $\vec{b}$ is in the image of the linear transformation $f(\vec{x}) = A\vec{x}$.

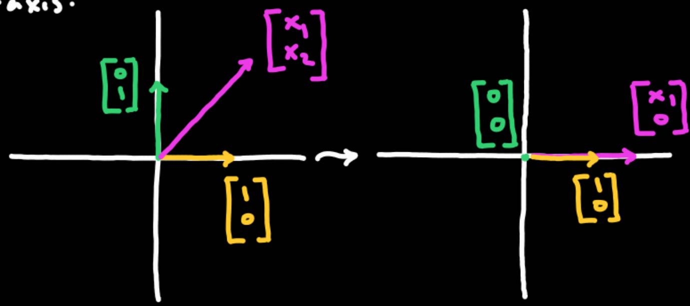
Since linear systems of equations can be visualized geometrically, and by the above these are equivalent to the image of linear transformations, we are able to visualize linear transformations geometrically.

Example: Describe what multiplication by the following matrices does geometrically.

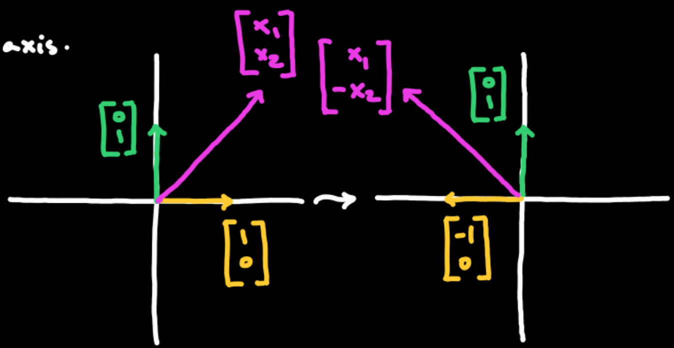
1) $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. Scaling by a factor of 2.



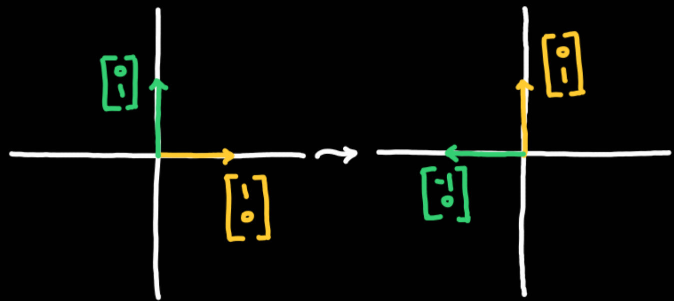
2) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Projection onto the x-axis.



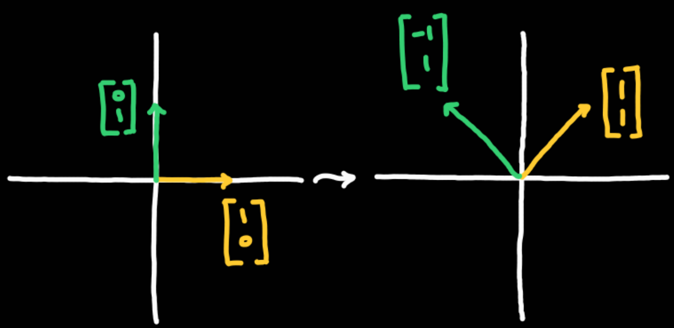
3) $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. Reflection about the y-axis.



4) $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Clockwise rotation by $\frac{\pi}{2}$.



5) $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Counter-clockwise rotation of $\frac{\pi}{4}$ and scaling by a factor of $\sqrt{2}$.



Observe that we can multiply the matrices of the rotation and the scaling to obtain this matrix:

$$\underbrace{\begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}}_{\text{scaling}} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{\text{rotation}} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{\text{rotation}} \underbrace{\begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}}_{\text{scaling}}.$$

6) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Shear towards the positive x-axis.