HW4 is up, due Thurs Oct 15.

Last time: Defined the notation

$$\text{“}A\vec{x} = \vec{b}\text{”}$$

as follows: Let A be $m \times n$ matrix

$$A = m\left\{\underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}}_{n}\right.$$

$$= \begin{pmatrix} | & | & & | \\ \vec{c}_1 & \vec{c}_2 & \cdots & \vec{c}_n \\ | & | & & | \end{pmatrix}$$

where $\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n \in \mathbb{R}^m$ are the column vectors of A. Then for any $n \times 1$ column vector $\vec{x} \in \mathbb{R}^n$ we define the $m \times 1$ column $A\vec{x} \in \mathbb{R}^m$:

$$A\vec{x} := x_1\vec{c}_1 + x_2\vec{c}_2 + \cdots + x_n\vec{c}_n \in \mathbb{R}^m$$

maybe the most important definition in the course ...

Alternatively, let

$$A = \begin{pmatrix} \text{—} \; \vec{r}_1^{\,T} \; \text{—} \\ \vdots \\ \text{—} \; \vec{r}_m^{\,T} \; \text{—} \end{pmatrix}$$

where $\vec{r}_1, \ldots, \vec{r}_m \in \mathbb{R}^n$ are the "row vectors" of the matrix A.

[Convention: $\vec{v}$ always denotes a column, $\vec{v}^{\,T}$ is always a row.]

Then for any $\vec{x} \in \mathbb{R}^n$ we have

$$A\vec{x} = \begin{pmatrix} \vec{r}_1^{\,T}\vec{x} \\ \vdots \\ \vec{r}_m^{\,T}\vec{x} \end{pmatrix}$$

where $\vec{r}_i^{\,T}\vec{x}$ = (row)(column) is the dot product $\vec{r}_i \cdot \vec{x}$.

Example: Express the following system as matrix equation $A\vec{x} = \vec{b}$:

$$\begin{cases} x + y + z = 5, \\ x + 2y + 4z = 3. \end{cases}$$

Solution:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

Compute $A\vec{x}$ in terms of <u>rows</u>:

$$\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + y + z$$

$$\begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + 2y + 4z$$

Hence

$$A\vec{x} = \begin{pmatrix} x+y+z \\ x+2y+4z \end{pmatrix}.$$

Or compute $A\vec{x}$ in terms of cols:

$$A\vec{x} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$= x\begin{pmatrix} 1 \\ 1 \end{pmatrix} + y\begin{pmatrix} 1 \\ 2 \end{pmatrix} + z\begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} x \\ x \end{pmatrix} + \begin{pmatrix} y \\ 2y \end{pmatrix} + \begin{pmatrix} z \\ 4z \end{pmatrix}$$

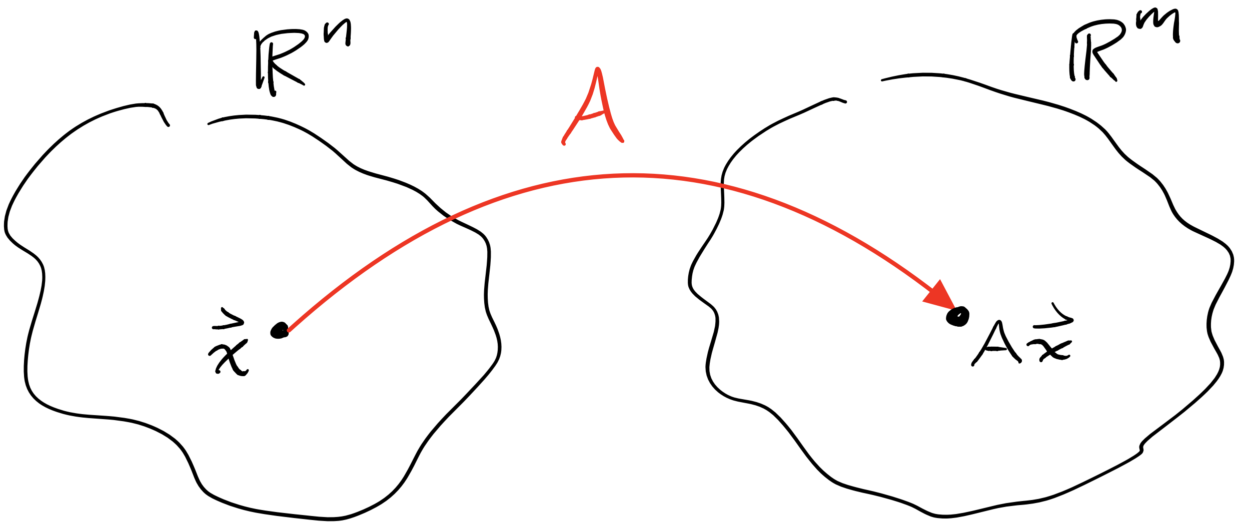
$$= \begin{pmatrix} x+y+z \\ x+2y+4z \end{pmatrix} \quad \text{SAME} \checkmark$$

Hence $A\vec{x} = \vec{b}$ says:

$$\begin{pmatrix} x+y+z \\ x+2y+4z \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \quad \checkmark$$

Today: A new point of view.

Given an $m \times n$ matrix A, we will think of this as a function from n-dimensional space to m-dimensional space:



This function has a very specific property called "linearity":

For all $\vec{u}, \vec{v} \in \mathbb{R}^n$ & $s, t \in \mathbb{R}$,

$$A(s\vec{u} + t\vec{v}) = sA\vec{u} + tA\vec{v}$$

"preserves linear combinations"

we can also express this property in two steps:

- $A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$

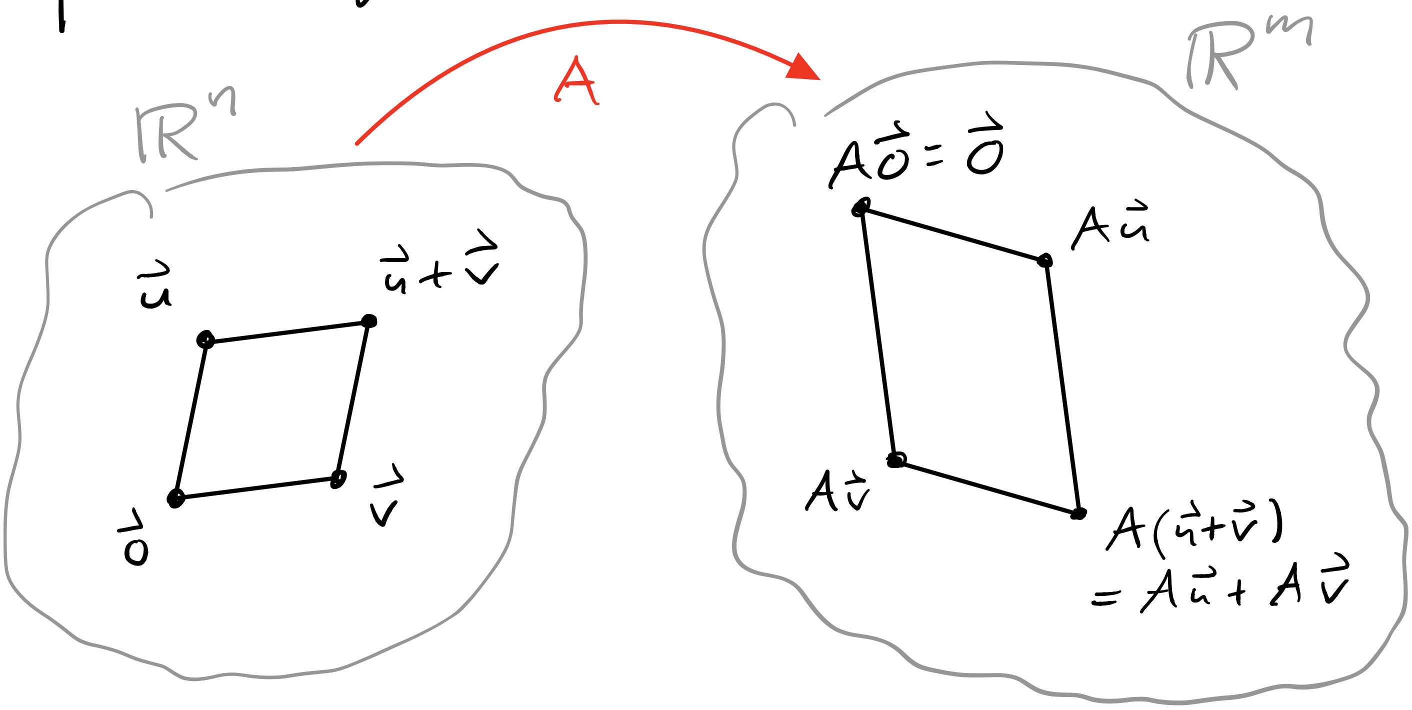
  "preserves vector addition"

- $A(t\vec{v}) = tA\vec{v}$

  "preserves scalar multiplication"

[You will prove this on HW 4.1.]

Geometric Meaning: A linear function sends parallelograms to parallelograms:



So what? Why do we care?

① Because every linear function comes from a matrix, and this tells us how to

"multiply matrices"

② Many basic geometric functions are linear, e.g. rotations, reflections, projections, ...

This is important in physics & computer graphics, but also (surprisingly) in probability & satistics. [Least Squares Approx.]

③ We can also think of linear functions between exotic kinds of vector spaces $f: U \to V$, e.g., derivatives & integrals,

$$\frac{d}{dx}\left(af(x)+bg(x)\right) = a\frac{d}{dx}f(x) + b\frac{d}{dx}g(x)$$

$$\int (af(x)+bg(x))\,dx = a\int f(x)\,dx + b\int g(x)\,dx.$$

① Linear functions $\mathbb{R}^n \to \mathbb{R}^m$ = $m\times n$ matrices.

We've already seen that $m\times n$ matrix A determines a linear function $A: \mathbb{R}^n \to \mathbb{R}^m$.

Conversely, let $f: \mathbb{R}^n \to \mathbb{R}^m$ be any function satisfying

$$f(s\vec{u}+t\vec{v}) = s f(\vec{u}) + t f(\vec{v}).$$

Goal: Find some $m\times n$ matrix A such that $f(\vec{x}) = A\vec{x}$ for all $\vec{x}$.

Solution: Let $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n \in \mathbb{R}^n$ be the "standard basis vectors"

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \vec{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Applying f to these gives some vectors $f(\vec{e}_1), \ldots, f(\vec{e}_n) \in \mathbb{R}^m$.

Define the matrix

$$A = \underbrace{\begin{pmatrix} | & & | \\ f(\vec{e}_1) & \cdots & f(\vec{e}_n) \\ | & & | \end{pmatrix}}_{n} \Big\} m$$

Then one can check [but not right now] that

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

as desired. ✓

Example: Let $id : \mathbb{R}^n \to \mathbb{R}^n$ be the "identity function"

$$id(\vec{x}) = \vec{x}$$

The corresponding matrix is called the $n \times n$ "identity matrix":

$$I = \begin{pmatrix} | & | & & | \\ id(\vec{e}_1) & id(\vec{e}_2) & \cdots & id(\vec{e}_n) \\ | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \cdots & \vec{e}_n \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & & 0 \\ 0 & 1 & 0 & & \vdots \\ \vdots & 0 & 1 & & \vdots \\ \vdots & \vdots & \vdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

Check: $I\vec{x} = \vec{x}$ for all $\vec{x}$.

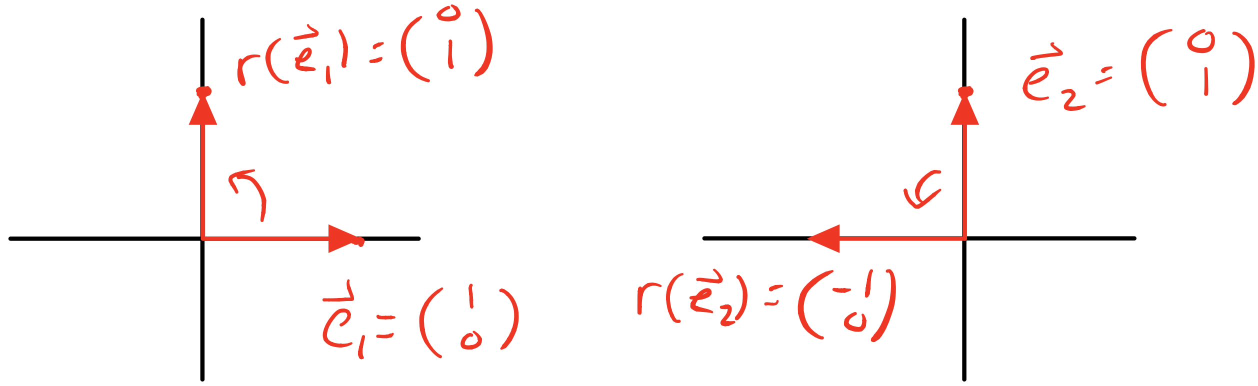
Example: Let $r: \mathbb{R}^2 \to \mathbb{R}^2$ be the function that rotates by $90°$ counterclockwise, which is a <u>linear</u> function (sends parallelograms to parallelograms). Find the 2×2 matrix $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that

$$r\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

How? We apply the function r to the standard basis vectors

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \;\&\; \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$



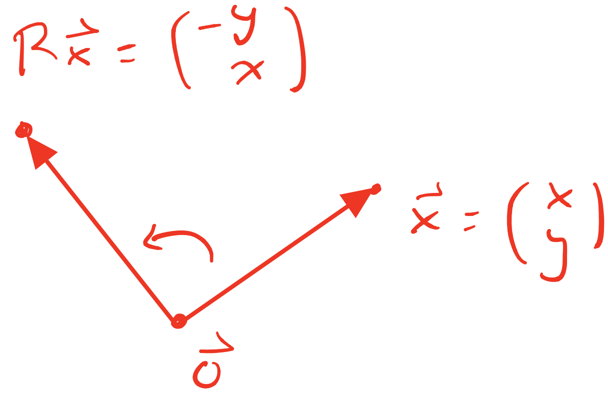
So now we know the matrix. Let's call it R:

$$R = \left(r(\vec{e}_1) \;\middle|\; r(\vec{e}_2) \right) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Conclusion: To rotate a general vector $\vec{x} = (x, y)$ multiply by R:

$$\begin{aligned} r(\vec{x}) &= R\vec{x} \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= x \begin{pmatrix} 0 \\ 1 \end{pmatrix} + y \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} \end{aligned}$$

Picture :



What if we rotate twice ?

$$r^2(\vec{x}) = (r \circ r)(\vec{x}) = r(r(\vec{x}))$$

"rotate by 180°"

To get the matrix :

$$r^2\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \& \quad r^2\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

We call this matrix

$$\text{"}R^2\text{"} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

This is an example of matrix multiplication !

$$R \cdot R = R^2$$

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

rotate twice by 90° c.c.w. = rotate by 180°

What about R^3 & R^4 ?

One can check that

$$R^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \text{rotate by } 90° \text{ \underline{clockwise}}$$

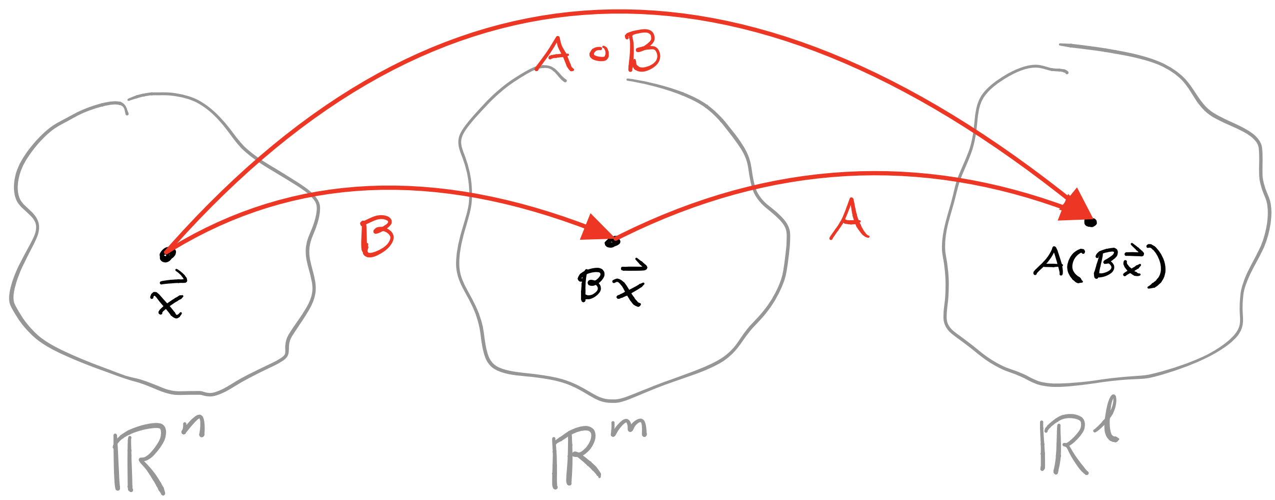
$$R^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \text{do nothing}$$

$= \underline{I}$ the 2×2 identity

How does matrix multiplication work in general ?

Let A be $\ell \times m$ matrix
B be $m \times n$ matrix.

Think of these as linear functions:



The composite function

$$A \circ B : \mathbb{R}^n \to \mathbb{R}^\ell$$

is also linear, so it corresponds to some $\ell \times n$ matrix. We call this matrix AB "A times B".

In other words, we define the matrix AB so that

$$A(B\vec{x}) = (AB)\vec{x}$$

[That seems reasonable.]

How to compute it?

Recall: For standard basis

$\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n \in \mathbb{R}^n$, we have

$$B\vec{e}_j = j\text{th column of } B.$$

$$(AB)\vec{e}_j = j\text{th column of } AB.$$

Hence

$$\begin{aligned} j\text{th col } AB &= (AB)\vec{e}_j \\ &= A(B\vec{e}_j) \\ &= A(j\text{th col of } B). \end{aligned}$$

That's it!

Example: $A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 2 & 0 \end{pmatrix}$

Compute AB & BA.

AB :

1st col AB = A (1st col B)

$$= \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1+0+4 \\ 1+0+2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

2nd col AB = A (2nd col B)

$$= \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1+0+0 \\ -1+1+0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Hence

$$AB = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ 3 & 0 \end{pmatrix}$$

Next BA :

1st col BA = B (1st col A)

$$= \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 - 1 \\ 0 + 1 \\ 2 + 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

2nd col BA = B (2nd col A)

$$= \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 - 1 \\ 0 + 1 \\ 0 + 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

3rd col BA = B (3rd col A)

$$= \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 - 1 \\ 0 + 1 \\ 4 + 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}.$$

Hence, $BA = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & 1 \\ 2 & 0 & 4 \end{pmatrix}.$

Important Observation:

$$AB \neq BA$$

$$\begin{pmatrix} 5 & -1 \\ 3 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & 1 \\ 2 & 0 & 4 \end{pmatrix}$$

These matrices don't even have the same shape!