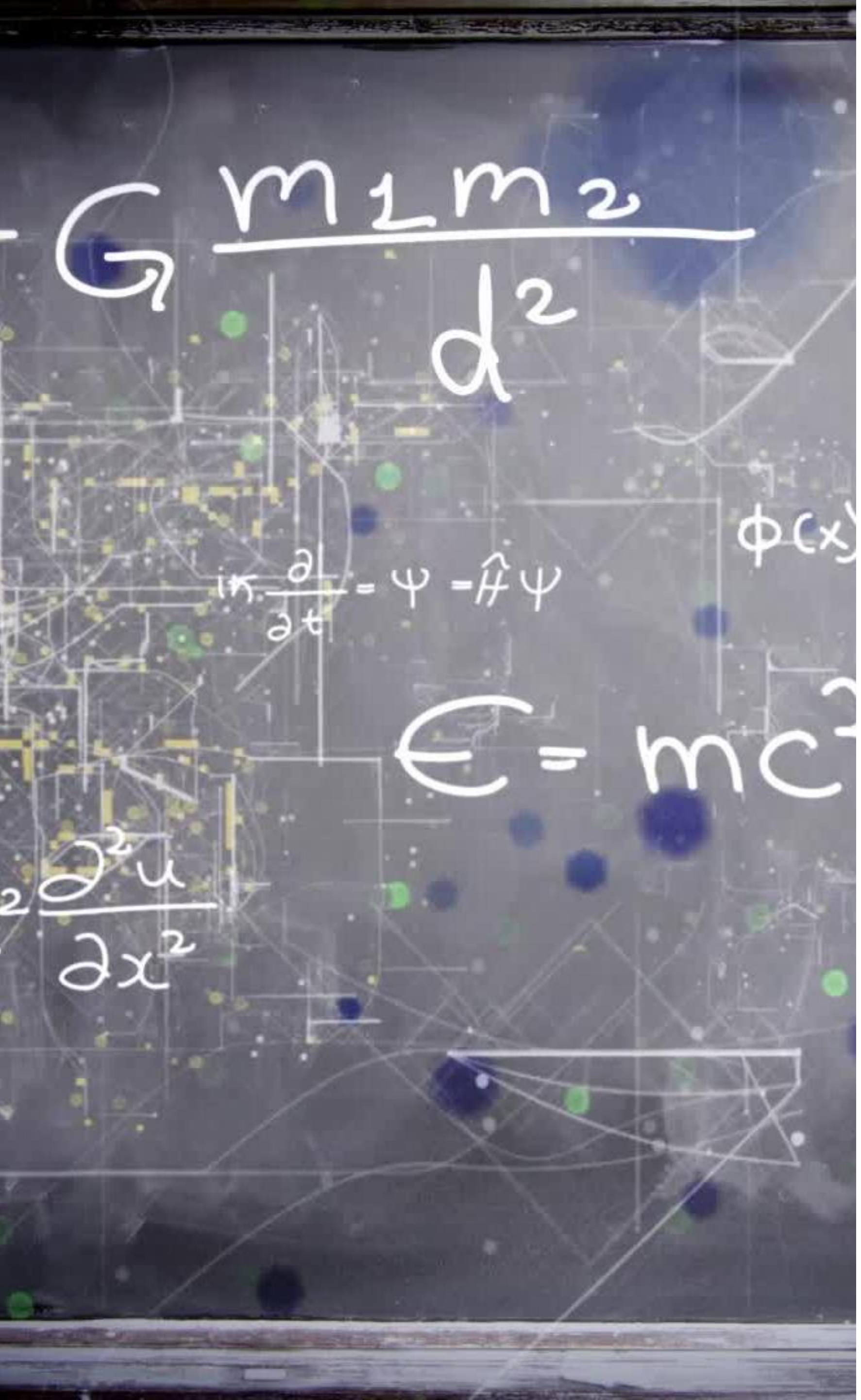
# Linear Algebra

Presented by Issa Ayoub



# Outline

- Motivation
- System of Sentences
- Terms and Definitions
- System of Equations
- Matrix Multiplication
- LU Factorization
- Vector and Linear Transformation
- Identity and Inverse Matrices
- Special kind of Matrices and Vectors
- Eigendecomposition
- Principle Component Analysis
- Sources

# Motivation

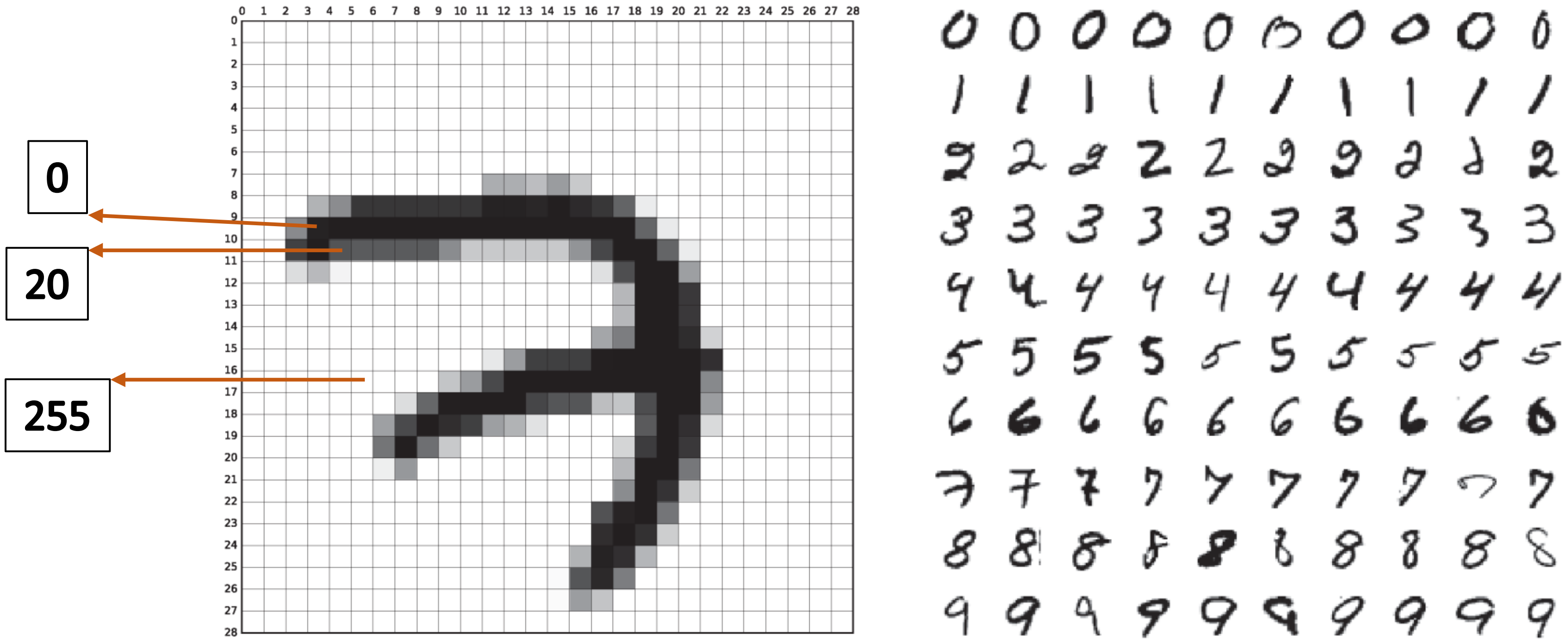
1. Data representation
    a. Iris dataset (example on tabular data)

| | sepal length (cm) | sepal width (cm) | petal length (cm) | petal width (cm) |
|---|---|---|---|---|
| 0 | 5.1 | 3.5 | 1.4 | 0.2 |
| 1 | 4.9 | 3.0 | 1.4 | 0.2 |
| 2 | 4.7 | 3.2 | 1.3 | 0.2 |
| 3 | 4.6 | 3.1 | 1.5 | 0.2 |
| 4 | 5.0 | 3.6 | 1.4 | 0.2 |

$$\begin{bmatrix} 5.1 & 3.5 & 1.4 & 0.2 \\ 4.9 & 3.0 & 1.4 & 0.2 \\ 4.7 & 3.2 & 1.3 & 0.2 \\ 4.6 & 3.1 & 1.5 & 0.2 \\ 5.0 & 3.6 & 1.4 & 0.2 \end{bmatrix}$$

# Motivation (CONT'D)

1. Data representation
    b. Mnist dataset (example on images)




(a) MNIST sample belonging to the digit '7'.

(b) 100 samples from the MNIST training set.

# Motivation (CONT'D)

1. Data representation
    c. Term Frequency-Inverse Document Frequency TF-IDF / word embeddings (examples on text)

*Assume 5 docs in total*

*In antiquity, Italy was the homeland of the Romans and the metropole of the Roman Empire's provinces. Rome was founded as*
*became a re*
*Roman mor*
*favor of a go*
*the People.*

*In antiquity, Italy was the homeland of the Romans and the metropole of the Roman Empire's provinces. Rome was founded as a kingdom in 753 BC and became a republic in 509 BC, when the Roman monarchy was overthrown in favor of a government of the Senate and the People.*

$$\begin{bmatrix} \mathrm{kw1} & \mathrm{kw2} & \mathrm{kw3} & \mathrm{kw4} & \mathrm{kw5} & \mathrm{kw6} & \mathrm{kw7} & \mathrm{kw8} & \mathrm{kw9} & \mathrm{kw10} \\ 0.1 & 0.0 & 0.2 & 0.0 & 0.3 & 0.4 & 0.0 & 0.2 & 0.1 & 0.0 \\ 0.0 & 0.3 & 0.0 & 0.2 & 0.1 & 0.0 & 0.3 & 0.4 & 0.2 & 0.1 \\ 0.2 & 0.1 & 0.0 & 0.3 & 0.0 & 0.1 & 0.0 & 0.2 & 0.4 & 0.3 \\ 0.3 & 0.0 & 0.1 & 0.0 & 0.2 & 0.3 & 0.4 & 0.0 & 0.1 & 0.2 \\ 0.0 & 0.2 & 0.3 & 0.1 & 0.0 & 0.2 & 0.1 & 0.3 & 0.0 & 0.4 \end{bmatrix}$$

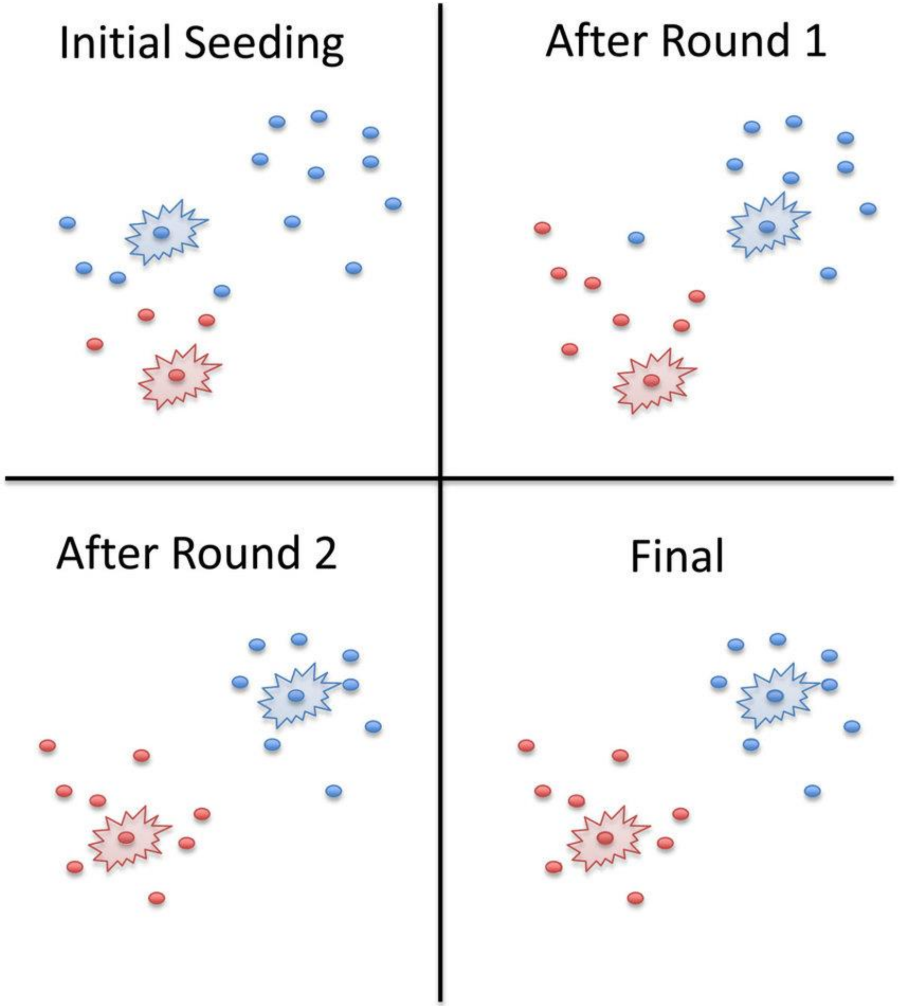
# Motivation (CONT'D)

2. Distance computation
   a. Euclidian distance (application: k-means).

| Point ID | X-coordinate | Y-coordinate |
|---|---|---|
| P1 | 2.3 | 4.2 |
| P2 | 3.1 | 6.8 |
| P3 | 5.5 | 1.9 |
| P4 | 7.0 | 4.1 |
| P5 | 4.2 | 8.5 |
| P6 | 1.7 | 2.6 |
| P7 | 6.3 | 3.5 |
| P8 | 8.1 | 6.9 |

$C_1$ : center 1

$C_2$ : center 2



# System of Sentences

**System 1**


The dog is **black**


The cat is **orange**

**Complete**

**Non-singular**

**System 2**


The dog is **black**


The dog is **black**

**Redundant**

**Singular**

**System 3**


The dog is **black**


The dog is **black**

**Contradictory**

**Singular**

# System of Sentences (CONT'D)

**System 1**

 The dog is **black**

 The cat is **orange**

 The bird is **red**

**Complete**

**Non-singular**

**System 2**

 The dog is black

 The dog is black

 The bird is **red**

**Redundant**

**Singular**

**System 3**

 The dog is **black**

 The dog is **black**

 The dog is **black**

**Redundant**

**Singular**

**System 4**

 The dog is **black**

 The dog is **white**

 The bird is **red**

**Contradictory**

**Singular**

# System of Sentences (CONT'D)
## Quiz: System of sentences

Given this system:

- Between the dog, the cat, and the bird, one I red.
- Between the dog and the cat, one is orange.
- The dog is black.

Problem 1:

What color is the bird?

Problem 2:

Is this system singular or non-singular?

Think about the following system of equations

$$\begin{aligned} x + y + z &= 6 \\ x + y &= 3 \\ x &= 1 \end{aligned}$$

# System of Sentences (CONT'D)
## System of Sentences vs. System of Equations

Compared to system of sentences:

$$x_1 = 5$$

$$x_2 = 9$$

In more formal terms:

$$4x_1 \ - 5x_2 = -13$$

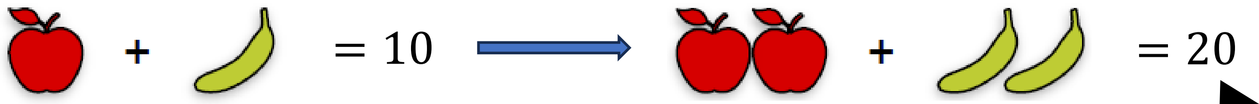
$$-2x_1 + 3x_2 = 9$$

$$Ax = b = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -13 \\ 9 \end{bmatrix}$$

# System of Sentences (CONT’D)

## Contradictory systems

- Day 1: You bought an apple and a banana, and they cost $10.




- Day 2: You bought two apples and two bananas, and they cost $24.


**Contradiction!**

- This is equivalent to:

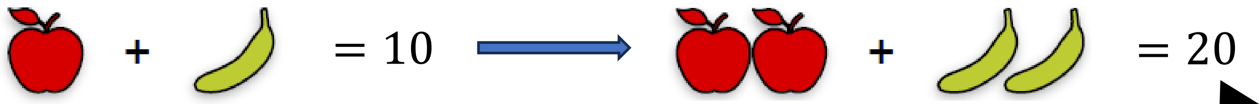
$$
\begin{aligned}
x + y &= 10 \\
2x + 2y &= 25
\end{aligned}
$$

**Contradictory Linear System!**

# System of Sentences (CONT'D)

## Redundant systems

- Day 1: You bought an apple and a banana, and they cost $10.




- Day 2: You bought two apples and two bananas, and they cost $20.


Infinite solutions!

- This is equivalent to:

$$x + y = 10$$
$$2x + 2y = 20$$

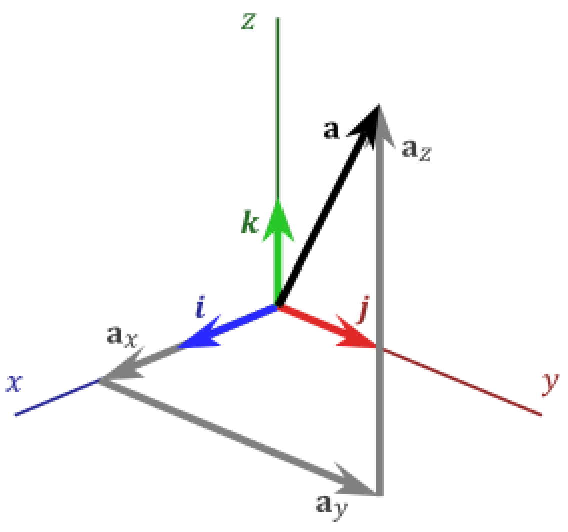
Infinite solutions for the linear system

# Terms and Definitions

- ***Scalars***:
  - A scalar is just a number, in contrast to other objects using in linear algebra such as arrays. Scalars are defined using the following notation expression:
  - "Let $\boldsymbol{s} \in \mathbb{R}$ be the slope of the line". This is how to define a *real valued number*.
  - To define *natural numbers*: "Let $\boldsymbol{s} \in \mathbb{N}$".
- ***Vectors***:
  - a vector is an array of numbers. Numbers are arranged in order. And numbers are identified by their index. Vectors are denoted using lower case names: $\boldsymbol{x}$
  - The first element in $\boldsymbol{x}$ is denoted as $x_1$ and the second element is $x_2$ $\Rightarrow$ italic typeface with a subscript.
  - Vectors have types too defined after their elements. If $\boldsymbol{x}$ contains 2 real variables => $\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{2}}$. More formally, $\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}$.
  - Vectors are identified by column vectors (by default) => $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

# Terms and Definitions (CONT'D)

- ***Vectors (CONT'D)***:
    - Vectors can be identified as points in space, with each element giving the coordinates along a different axis.



# Terms and Definitions (CONT’D)

- ***Matrices***:
  - A matrix is a 2-D array of numbers.
  - Each element is identified by two indices. E.g.: $A_{(i,j)}\ or\ a_{ij}$
  - A matrix containing only one column is called a **column vector,** and a matrix containing only one row is a **row vector**.

*Column matrix/column vector:*

$$x = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

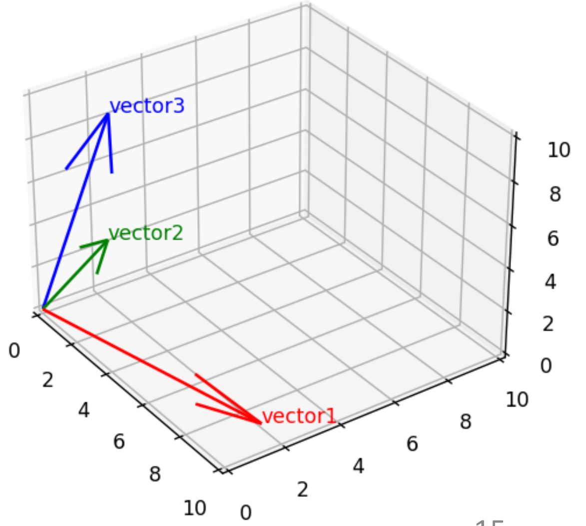
*Row matrix/row vector:*

$$x = \begin{bmatrix} a & b & c \end{bmatrix}$$

  - Matrices as well can be seen as vectors in n-D space. Imagine the following matrix:

$$\begin{bmatrix} 9 & 2 & 1 \\ 2 & 0 & 2 \\ 0 & 2 & 9 \end{bmatrix}$$

*vector1* *vector2* *vector3*



# System of equations

## What is a System of Linear Equations

- A system of linear equation consists of ***m*** linear equations and ***n*** unknown variables $x_1, x_2, \dots, x_n$.

- **Problem**: Find a list of ***n*** numbers $s_1, s_2, \dots, s_n$ that satisfy the system of linear equations ⇒ if we substitute $x_1, x_2, \dots, x_n$ by $s_1, s_2, \dots, s_n$, the left hand side of the $i^{th}$ equation will equal $b_i$ for $i \in [1, \dots, m]$.

- We call $s_1, s_2, \dots, s_n$ "a solution" to the system of equations.
  - "a solution" because we could have more than one solution.
  - A solutions are called **solution set**.

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
$$

# System of equations (CONT'D)

## What is a System of Linear Equations

- Below is a linear system of $m = 3$ equations and $n = 4$ unknowns.

$$\begin{aligned} 2x_1 + 3x_2 - x_3 + 4x_4 &= 5 \\ -3x_1 + 4x_2 + 7x_3 - 2x_4 &= 1 \\ 5x_1 - 2x_2 + x_3 + 3x_4 &= 10 \end{aligned}$$

- <u>Quiz</u>: is (1, -1, 2) a solution to the following system of linear equations:
  - What about (1, 2, -4)

$$\begin{aligned} 2x_1 + 2x_2 + x_3 &= 2 \\ x_1 + 3x_2 - x_3 &= 11. \end{aligned}$$

# System of equations (CONT'D)

## What is a System of Linear Equations

- **Definition**: A linear system may not have a solution at all. If this is the case, we say that the linear system is ***inconsistent***.

**INCONSISTENT ⇔ NO SOLUTION**

***What about the term singular?***

- **Definition**: A linear system is called ***consistent*** if it has at least one solution.

**CONSISTENT ⇔ # SOLUTIONS >= 1**

- **Note**: a ***consistent*** linear system cannot have 4 or 5 solution

back

# System of equations (CONT'D)

## What is a System of Linear Equations

- Example 1: Show that the following linear system doesn't have a solution:

$$\begin{aligned} -x_1 + x_2 &= 3 \\ x_1 - x_2 &= 1. \end{aligned}$$

- Solution: if we add both equations we can: $0 = 4$

**Contradictory**

**Singular**

Recall from system of sentences

# System of equations (CONT'D)
## Matrices

- Matrices are used to develop a systematic method to solve the linear system and to study the properties of the solution set of a linear system.
- We can derive three matrices for any linear system:
  - The coefficient matrix
  - The output column vector
  - The augmented matrix
- For instance, consider the following linear system:
  - The three matrices are:

$$\begin{aligned} 2x_1 + 3x_2 - x_3 + 4x_4 &= 5 \\ -3x_1 + 4x_2 + 7x_3 - 2x_4 &= 1 \\ 5x_1 - 2x_2 + x_3 + 3x_4 &= 10 \end{aligned}$$

$$A = \begin{bmatrix} 2 & 3 & -1 & 4 \\ -3 & 4 & 7 & -2 \\ 5 & -2 & 1 & 3 \end{bmatrix} \qquad b = \begin{bmatrix} 5 \\ 1 \\ 10 \end{bmatrix} \qquad [A \quad b] = \left[\begin{array}{cccc|c} 2 & 3 & -1 & 4 & 5 \\ -3 & 4 & 7 & -2 & 1 \\ 5 & -2 & 1 & 3 & 10 \end{array}\right]$$

# System of equations (CONT'D)

## Solving Linear Systems

- Common linear algebra operations:
    1. Add some value to both sides of the equation
    2. Multiply both sides by the same scalar.
    - E.g. $\boxed{2x = 8 - 2x}$ $\Rightarrow$ we can add $2x$ on both sides of the equation $\Rightarrow 4x = 8$ (step 1),
    - Then divide by ¼ on both sides (step 2) $\Rightarrow x = 2$.
- Similar operations can be applied to linear systems. There are only 3 basic operations named ***elementary operations***, that can performed:
    - ***Interchange two equations. (1)***
    - ***Multiply an equation by a non zero constant. (2)***
    - ***Add a multiple of one equation to another. (3)***

*These operations do not alter the solution set.*

Question: how can we add a constant to both sides of an equation within the context of linear systems.

# System of equations (CONT'D)

## Solving Linear Systems / Row Echelon Form

- E.g.:
  - ***Interchange two equations. (1)***

$$\begin{aligned} 2x_1 + 3x_2 - x_3 + 4x_4 &= 5 \\ -3x_1 + 4x_2 + 7x_3 - 2x_4 &= 1 \end{aligned} \equiv \begin{aligned} -3x_1 + 4x_2 + 7x_3 - 2x_4 &= 1 \\ 2x_1 + 3x_2 - x_3 + 4x_4 &= 5 \end{aligned}$$

  - ***Multiply an equation by a non-zero constant. (2)***

$$-3x_1 + 4x_2 + 7x_3 - 2x_4 = 1 \equiv -6x_1 + 8x_2 + 14x_3 - 4x_4 = 2$$

  - ***Add a multiple of one equation to another. (3)***

$$\begin{aligned} 2x_1 + 3x_2 - x_3 + 4x_4 &= 5 \\ -3x_1 + 4x_2 + 7x_3 - 2x_4 &= 1 \end{aligned} \equiv \begin{aligned} 2x_1 + 3x_2 - x_3 + 4x_4 &= 5 \\ -x_1 + 7x_2 + 6x_3 + 2x_4 &= 6 \end{aligned}$$

# System of equations (CONT'D)

## Solving Linear Systems / Row Echelon Form

- When those operations are applied to the coefficient matrix or augmented matrix, we call them ***elementary row operations***.
- The process of simplifying the linear system using these operations is called ***row reduction***.
- The goal with row reducing is to transform the original linear system into one having a **triangular structure** and then performing **back substitution** to solve the system. This is best explained with an example:

$$\begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$x_1 - 2x_3 = -4 \Rightarrow x_1 = 2x_3 - 4 \Rightarrow \boxed{x_1 = -2}$$

$$x_2 - x_3 = 0 \Rightarrow \boxed{x_2 = 1}$$

$$\boxed{x_3 = 1}$$

# System of equations (CONT'D)

## Solving Linear Systems / Row Echelon Form

- E.g. 2: Assume we have the following linear system:

The augmented matrix is:

$$\begin{aligned} -3x_1 + 2x_2 + 4x_3 &= 12 \\ x_1 - 2x_3 &= -4 \\ 2x_1 - 3x_2 + 4x_3 &= -3 \end{aligned} \longrightarrow \begin{bmatrix} -3 & 2 & 4 & 12 \\ 1 & 0 & -2 & -4 \\ 2 & -3 & 4 & -3 \end{bmatrix}$$

Interchange Row 1 (R1) and Row 2 (R2)

$$\begin{bmatrix} -3 & 2 & 4 & 12 \\ 1 & 0 & -2 & -4 \\ 2 & -3 & 4 & -3 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 0 & -2 & -4 \\ -3 & 2 & 4 & 12 \\ 2 & -3 & 4 & -3 \end{bmatrix}$$

This first operation will simplify the next step. Add $3R_1$ to $R_2$.

$$\begin{bmatrix} 1 & 0 & -2 & -4 \\ -3 & 2 & 4 & 12 \\ 2 & -3 & 4 & -3 \end{bmatrix} \xrightarrow{3R_1 + R_2} \begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 2 & -2 & 0 \\ 2 & -3 & 4 & -3 \end{bmatrix}$$

# System of equations (CONT'D)

## Solving Linear Systems / Row Echelon Form

Add -2$R_1$ to $R_3$.

$$\begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 2 & -2 & 0 \\ 2 & -3 & 4 & -3 \end{bmatrix} \xrightarrow{-2R_1+R_3} \begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 2 & -2 & 0 \\ 0 & -3 & 8 & 5 \end{bmatrix}$$

Multiply $R_2$ by ½ .

$$\begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 2 & -2 & 0 \\ 0 & -3 & 8 & 5 \end{bmatrix} \xrightarrow{\frac{1}{2}R_2} \begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 1 & -1 & 0 \\ 0 & -3 & 8 & 5 \end{bmatrix}$$

Add 3$R_2$ to $R_3$ .

$$\begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 1 & -1 & 0 \\ 0 & -3 & 8 & 5 \end{bmatrix} \xrightarrow{3R_2+R_3} \begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 5 & 5 \end{bmatrix}$$

Multiply $R_3$ by 1/5 .

$$\begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 5 & 5 \end{bmatrix} \xrightarrow{\frac{1}{5}R_3} \boxed{\begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}}$$

The row reduced augmented matrix is in triangular form.

# System of equations (CONT'D)

## Solving Linear Systems / Row Echelon Form

Therefore, we have moved the original augmented matrix

$$\begin{bmatrix} -3 & 2 & 4 & 12 \\ 1 & 0 & -2 & -4 \\ 2 & -3 & 4 & -3 \end{bmatrix} \xrightarrow{\textit{Forward Elimination}} \begin{bmatrix} 1 & 0 & -2 & -4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

To the row echelon form augmented matrix, which is in triangular form.

Now, the linear system can now be solved using back substitution. This is the same augmented matrix we saw at slide 21.

Its equivalent linear system is:

$$\begin{aligned} -3x_1 + 2x_2 + 4x_3 &= 12 \\ x_1 - 2x_3 &= -4 \\ 2x_1 - 3x_2 + 4x_3 &= -3 \end{aligned}$$

Its equivalent linear system is:

$$\begin{aligned} x_1 - 2x_3 &= -4 \\ x_2 - x_3 &= 0 \\ x_3 &= 1. \end{aligned}$$

Those are effectively the same linear system with the same solution set.

# System of equations (CONT'D)

## Solving Linear Systems / Row Echelon Form

- What does it mean for the linear system to be inconsistent?
- E.g., consider the following system of linear equations:

$$\begin{aligned} x_1 + 2x_3 &= 1 \\ x_2 + x_3 &= 0 \\ 2x_1 + 4x_3 &= 1 \end{aligned} \xrightarrow{\textit{Its augmented matrix is:}} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 4 & 1 \end{bmatrix}$$

- After applying row reduction, we obtain the following upper triangular matrix: $\begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$
- However, the last row corresponds to the following equation:

$$0x_1 + 0x_2 + 0x_3 = -1$$

- Therefore, there are no numbers $x_1, x_2$, and $x_3$ that satisfy this equation, therefore, the linear system is ***inconsistent***.

back

# System of equations (CONT'D)

## Solving Linear Systems / Row Echelon Form

- Therefore, as a rule of thumb, when the last row in the row reduced augmented matrix is of the form: $\boldsymbol{[0, 0, \dots, c]}$, the former linear system is inconsistent.
    - We call this type of row inconsistent row.
    - However: $[0, 1, \dots, 0]$ is perfectly valid. This is equivalent to: $x_2 = 0$ which is perfectly valid. (Slide 16)

# System of equations (CONT'D)

## Solving Linear Systems / Row Echelon Form

- Consider the following linear system and its augmented matrix:

$$\begin{aligned} x_1 + 5x_2 - 2x_4 - x_5 + 7x_6 &= -4 \\ 2x_2 - 2x_3 + 3x_6 &= 0 \\ -9x_4 - x_5 + x_6 &= -1 \\ 5x_5 + x_6 &= 5 \\ 0 &= 0 \end{aligned}$$

$$\begin{bmatrix} 1 & 5 & 0 & -2 & -1 & 7 & -4 \\ 0 & 2 & -2 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -9 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 5 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

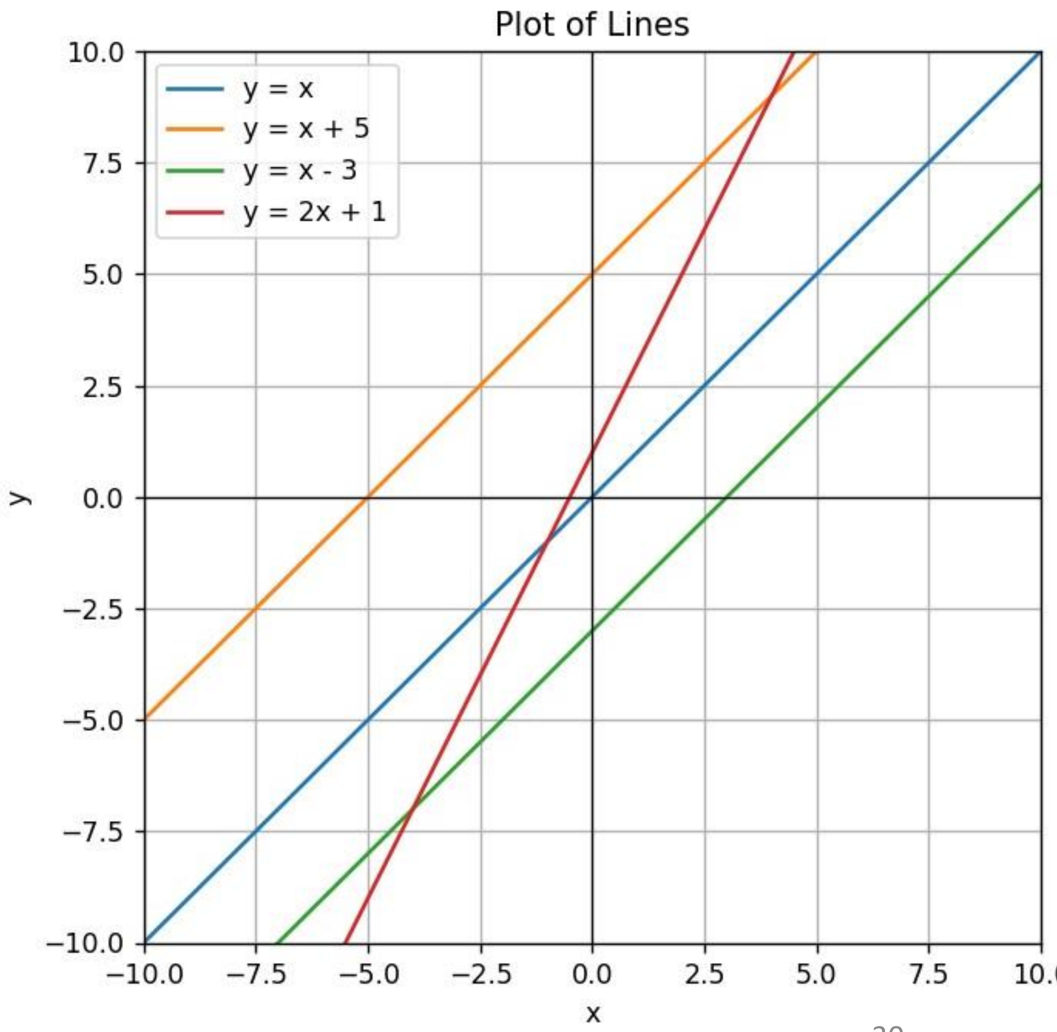
- The augmented matrix has the following ***Properties:***
  - ***Property1:*** All nonzero rows are above any rows of all zeros.
  - ***Property2:*** The leftmost nonzero entry of a row is to the right of the leftmost nonzero entry of the row above it.
- Any matrix satisfying properties P1 and P2 are said to be in **row echelon form**

back

# System of equations (CONT’D)

## Geometric Interpretation of the Solution Set

- We can look at a system by rows or by columns. The first approach concentrates on the separate equations (the rows).
- Consider the following lines of equations:
  - $y = x + 5$ (line 1)
  - $y = x - 3$ (line 2)
  - $y = x$ (line 3)
  - $y = 2x + 1$ (line 4)
  - How they differ?
- Changing the constant after the equal sign is equivalent to shifting the line up or down.
- Changing the slope of the line is equivalent to rotating the line by some angle.



# System of equations (CONT'D)

## Geometric Interpretation of the Solution Set

- If we consider the same equations, and arrange them into a linear system to find the solution set we get the following:
  - $y = x + 5 \quad \Rightarrow \quad \mathrm{y} - \mathrm{x} = 5$
  - $y = x - 3 \quad \Rightarrow \quad \mathrm{y} - \mathrm{x} = -3$
  - What does it mean to find a solution?
  - Can you find a solution?

$$[A \;\; b] = \begin{bmatrix} 1 & -1 & 5 \\ 1 & -1 & -3 \end{bmatrix}$$

- Therefore, when the linear system involves parallel lines, those are equivalent to contradictory linear equations.
- Therefore, two parallel lines will not intersect and the same applies to parallel planes. (lines 1, 2, and 3)
  - This is equivalent to contradictory equations. => the system of linear equations is **inconsistent or singular**. (Slide 16)

# System of equations (CONT’D)

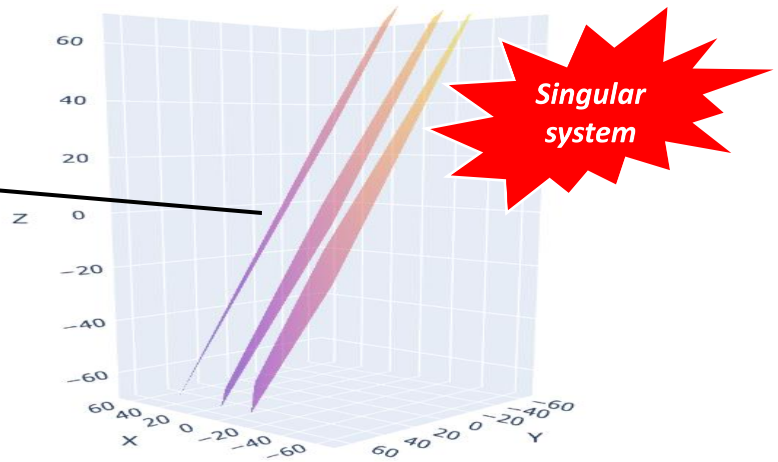
## Geometric Interpretation of the Solution Set

- Now, to all form of possible answers.
- Link to [webapp]
- Consider the following matrices:
  - ***Inconsistent results***:

**No solutions(contradictory)**

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & -20 \\ 1 & 1 & 1 & 30 \end{bmatrix}$$

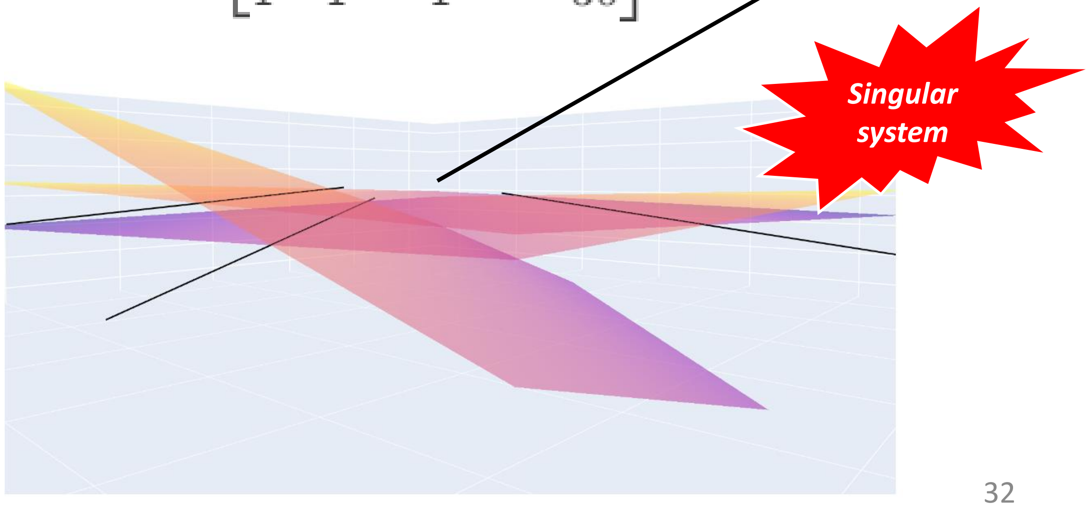
All planes are parallel




**No solutions (contradictory)**

$$A = \begin{bmatrix} 1 & 1 & 10 & 0 \\ 1 & 1 & -10 & 50 \\ 1 & 1 & 1 & -30 \end{bmatrix}$$

*Every two planes Intersecting in different line*



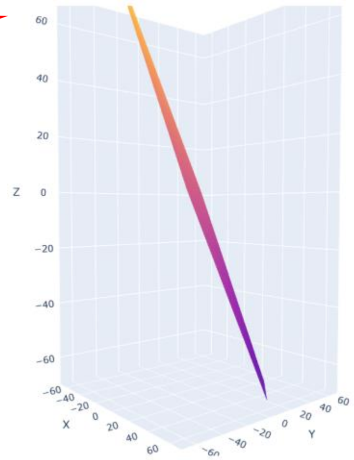
# System of equations (CONT'D)

## Geometric Interpretation of the Solution Set

- Link to webapp
- Consider the following matrices:
    - ***Inconsistent linear system (Infinite solutions):***

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

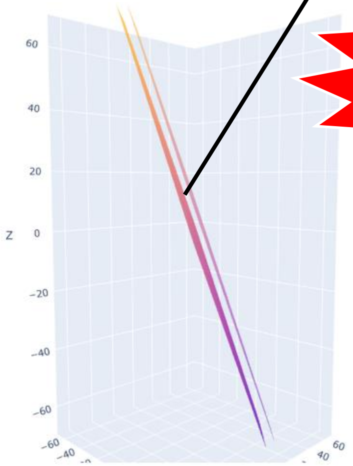
**Singular system**




$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 10 \end{bmatrix}$$

*Two of the planes Overlap (they both pass through the origin)*

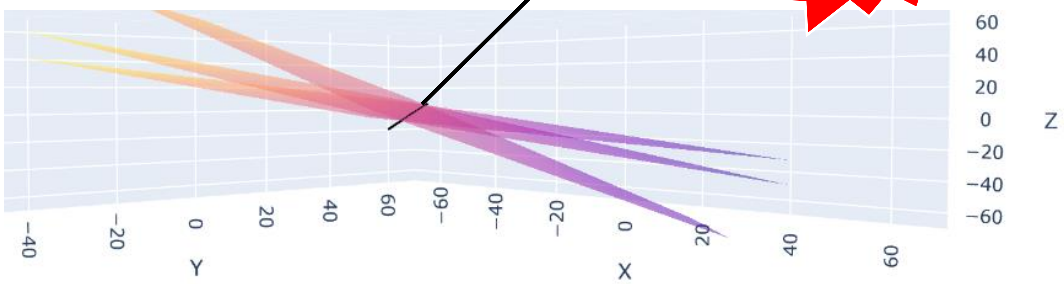
**Singular system**




$$A = \begin{bmatrix} 1 & 1 & 1 & 10 \\ 1 & 1 & 2 & 15 \\ 1 & 1 & 3 & 20 \end{bmatrix}$$

*Every point on the line of intersection is a solution*

**Singular system**



# System of equations (CONT’D)

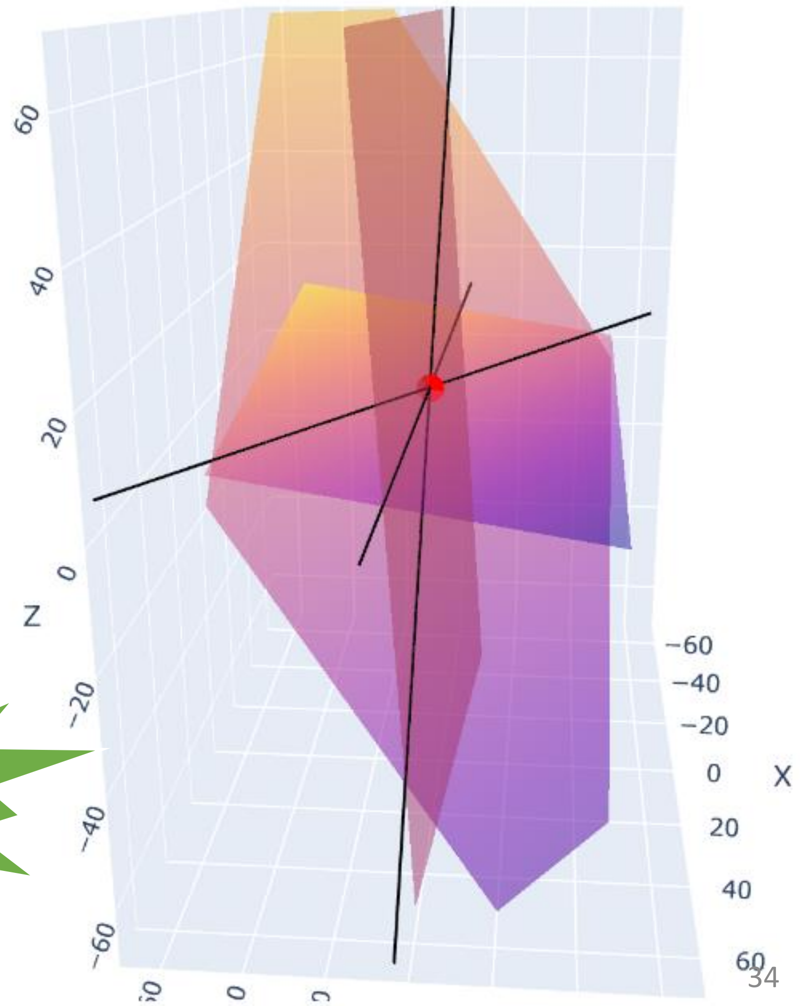
## Geometric Interpretation of the Solution Set

- Link to [webapp]
- Consider the following matrices:
  - ***Consistent linear system:***
  - The point of intersection lies in all 3 planes. And this single point is the only solution for this linear system.
  - Each plane of those is somehow has dimensions = 2.

***Unique solutions***

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 10 & 50 \\ 1 & 10 & 1 & -30 \end{bmatrix}$$

***Non-Singular system***



# System of equations (CONT'D)

## Geometric Interpretation of the Solution Set

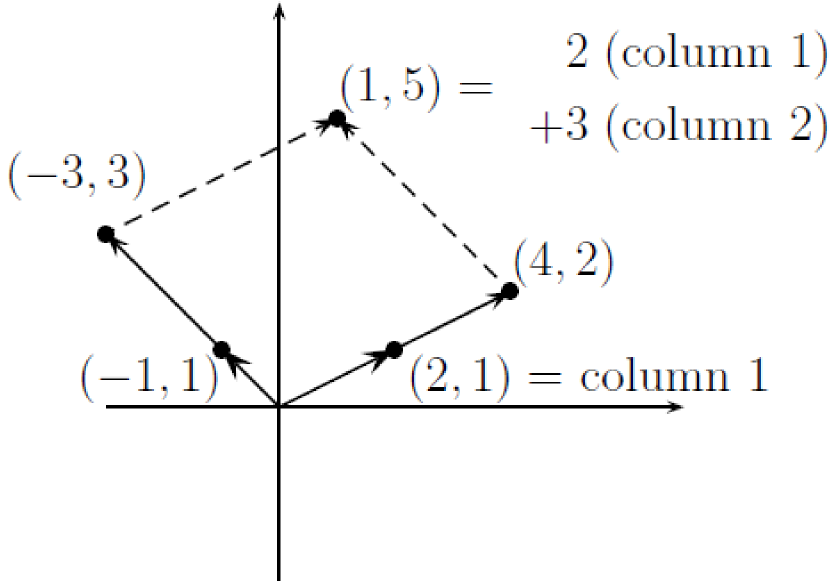
- **The second approach looks at the columns of the linear system.**

*Each equation describes a plane in three dimensions*

**Column form** $x\begin{bmatrix}2\\1\end{bmatrix}+y\begin{bmatrix}-1\\1\end{bmatrix}=\begin{bmatrix}1\\5\end{bmatrix}$

**Row form** $\boxed{2x-y=1}$ $\boxed{x+y=5}$

- The problem is to find the **combination of the column vectors** on the left side that produces the vector on the right side.
- The unknowns are the numbers x and y that multiply the column vector.



# System of equations (CONT'D)
## Geometric Interpretation of the Solution Set

- In more formal terms, given vectors $v_1, \dots, v_p$ and $b$, is $b$ a linear combination of $v_1, v_2, \dots, v_p$?
- For instance, if we are given the following vectors:

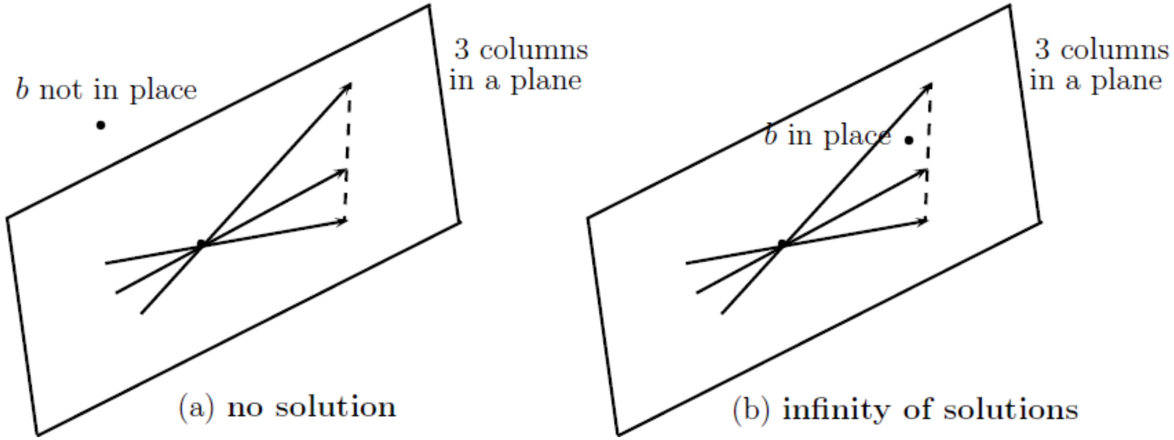
$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}$$

- Does there exist scalars $x_1, x_2, x_3$ such that: $x_1 v_1 + x_2 v_2 + x_3 v_3 = b$?
- The problem of determining if a given vector b is a linear combination of the vectors $v_1, v_2, \dots, v_p$ is equivalent to solving the linear system of equations with the augmented matrix.

$$[A \quad b] = [v_1 \quad v_2 \quad \dots \quad v_p \quad b]$$

# System of equations (CONT’D)
## Geometric Interpretation of the Solution Set

- The span of a set of vectors $\{x_1, x_2, \dots, x_n\}$ is the set of all vectors that can be expressed as a linear combination of $\{x_1, x_2, \dots, x_n\}$.
- If $\{x_1, x_2, \dots, x_n\}$ are $n$ linearly independent vectors where each $x_i \in \mathbb{R}^n$ , then $span(\{x_1, x_2, \dots, x_n\}) = \mathbb{R}^n$.
    - In other words, any vector $v \in \mathbb{R}^n$ can be written as a linear combination of $x_1$through $x_n$.
- The ***range*** (also called column space) of a matrix $A \in \mathbb{R}^{m \times n}$, denoted $\mathcal{R}(A)$ is the span of the columns of $A$.
- What happens to the column picture when the system is singular?




(a) no solution

(b) infinity of solutions

# System of equations (CONT’D)
## Linear dependence between rows/Cols

- Assume having the following linear system

*This comes down to*

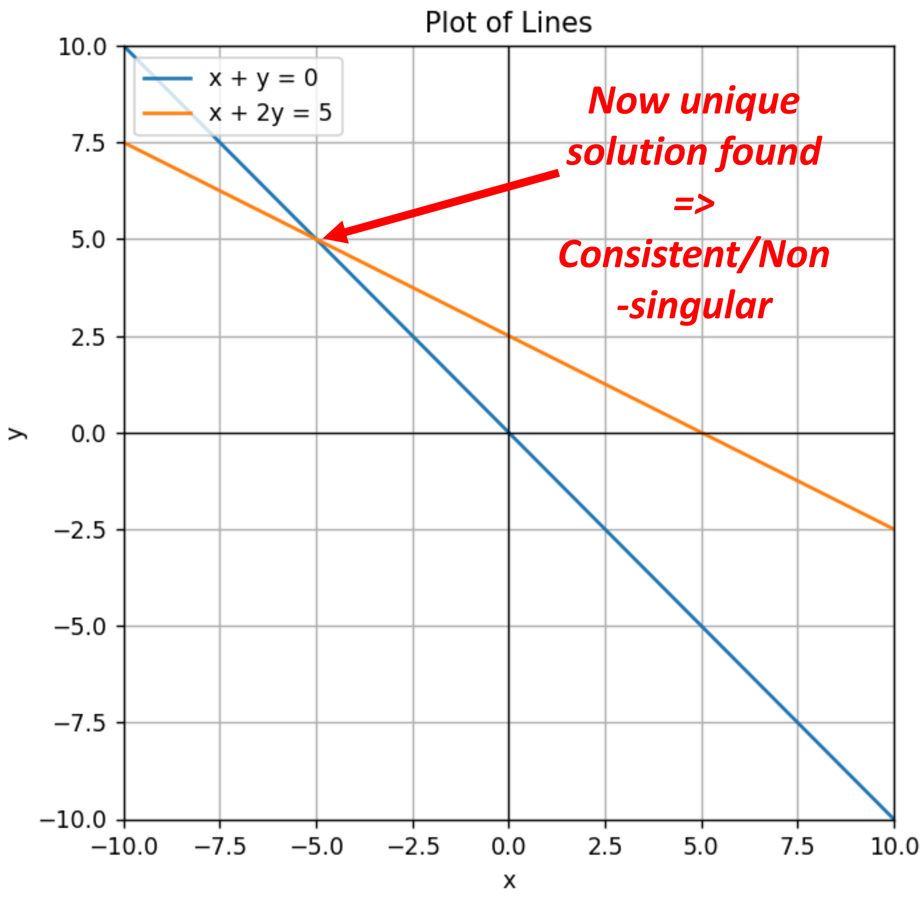
$$\begin{aligned} x + y &= 1 \\ 2x + 2y &= 5 \end{aligned} \quad \Rightarrow \quad \begin{aligned} x + y &= 1 \\ x + y &= 2.5 \end{aligned}$$

*When the coefficients of the variables are equal then the lines are parallel. The same applies in higher dimensional planes.*

- Now change the coefficient of x or y in one of the equations:

$$x + y = 1$$

$$x + 2y = 2.5$$



# System of equations (CONT'D)
## Linear dependence between rows/cols

- Therefore, if we look at the coefficient matrix:

$$\begin{aligned} x + y &= 1 \\ x + y &= 2.5 \end{aligned} \Rightarrow \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \Rightarrow$$

*Row 2 = 2 * row 1* $\Rightarrow$ *OR Col 1 = Col 2*

**Singular / Inconsistent**

- However, for the second linear system:

$$\begin{aligned} x + y &= 1 \\ x + 2y &= 2.5 \end{aligned} \Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow$$

*No dependence between rows and cols.*

**Non-singular / consistent**

# System of equations (CONT'D)
## Rank

- A set of vectors are said to be ***(linearly) independent*** if no vector can be expressed as a linear combination of other vectors within the same set.
- If any vector can be expressed as a linear combination of other vectors, then the vectors are said to be ***(Linearly) dependent***. In this case:

$$x_n = \sum_{i=1}^{n-1} \alpha_i x_i$$

  - For some scalar values $\alpha_1, \dots, \alpha_{n-1} \in \mathbb{R}$.
- For example, the following vectors are linearly dependent because $x_3 = -2x_1 + x_2$.

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad x_2 = \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix} \quad x_3 = \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}$$

# System of equations (CONT'D)
## Rank

- The ***column rank*** of a matrix is the size of the largest subset of columns that constitute a linearly independent set. This is simply the number of linearly independent columns in the coefficient matrix.
- Similarly, the ***row rank*** is the largest number of independent rows that constitute a linearly independent set.
- For any matrix, row rank is equal to column rank.
  - Both quantities are referred to collectively as the rank.
- For any matrix $A \in \mathbb{R}^{m\times n}$, $\mathrm{rank(A)} \leq \mathrm{min(m, n)}$, If $\mathrm{rank(A)} = \mathrm{min(m, n)}$, then A is said to be ***full rank.***
- For $A \in \mathbb{R}^{m\times n}$, $\mathrm{rank(A)} = \mathrm{rank(A}^T)$
- For $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times p}$, $\mathrm{rank(A}B) \leq \min(rank(A), rank(B))$.
- For $A, B \in \mathbb{R}^{m\times n}$, $\mathrm{rank(A} + B) \leq \mathrm{rank(A)} + \mathrm{rank}(B)$

# System of equations (CONT'D)

## Determinant

- Again, assume we have the following matrix:
- This matrix is singular if $[a, b] * k = [c, d] \Rightarrow$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$ak = c$$
$$bk = d$$

$$\frac{c}{a} = \frac{d}{b} = k$$

$$ad = bc$$

***determinant*** $\longrightarrow \boxed{ad - bc = 0}$

$\Rightarrow \boxed{determinant = ad \; - bc}$

a d - c b

# System of equations (CONT'D)
## Determinant

- Therefore,

$$if\ determinant = ad - cb = 0$$
$$\Rightarrow k \neq 0$$
$$\Rightarrow$$ ***There exist dependent rows/cols***
$$\Rightarrow$$ ***Matrix is Singular/Inconsistent***

***WHY?***

*Recall parallel planes!*

Note: when calculating the determinant, it is sufficient to look at the coefficient matrix rather than the augmented matrix.

*Recall:*
$$determinant = ad - bc$$

| | | | | |
|---|---|---|---|---|
| a | | - | | b |
| | d | | c | |

# System of equations (CONT'D)

## Determinant-Quiz

- Find the determinant of the following matrices

$$\begin{bmatrix} 5 & 1 \\ -1 & 3 \end{bmatrix} \qquad \det = 5.3 - 1.(-1) = 15 + 1 = 16$$

Non-singular / consistent

$$\begin{bmatrix} 2 & -1 \\ -6 & 3 \end{bmatrix} \qquad \det = 2.3 - (-1).(-6) = 6 - 6 = 0$$

Singular / Inconsistent

- Are those matrices Singular or Non-singular?

# System of equations (CONT’D)

## Reduced Row Echelon Form

- Although REF simplifies the problem of solving a linear system, we can further reduce the matrix into what is called ***reduced row echelon form (RREF)***.
- A matrix is in RREF if it satisfies properties 1 and 2 ([@slide 29]) and in addition satisfies the following ***properties***:
  - Property 3: The leading entry in each nonzero row is a 1.
  - Property 4: All the entries above (and below) a leading 1 are all zeros.
- A leading 1 in the RREF of a matrix is called a pivot.

*pivots*

$$\begin{bmatrix} \boxed{1} & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & \boxed{1} & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & \boxed{1} & 7 \end{bmatrix}$$

# System of equations (CONT'D)
## Reduced Row Echelon Form

- Example, use row reduction to transform the following matrix into RREF.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

- The first step is to make the top leftmost entry nonzero

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \xrightarrow{R_3 \leftrightarrow R_1} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

- Then create a leading 1 in the first row

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \xrightarrow{\frac{1}{3}R_1} \begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

- Create zeros under the newly created leading 1:

$$\begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \xrightarrow{-3R_1+R_2} \begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

# System of equations (CONT'D)
## Reduced Row Echelon Form

- Create a leading 1 in the second row

$$\begin{bmatrix}1 & -3 & 4 & -3 & 2 & 5\\0 & 2 & -4 & 4 & 2 & -6\\0 & 3 & -6 & 6 & 4 & -5\end{bmatrix} \xrightarrow{\frac{1}{2}R_2} \begin{bmatrix}1 & -3 & 4 & -3 & 2 & 5\\0 & 1 & -2 & 2 & 1 & -3\\0 & 3 & -6 & 6 & 4 & -5\end{bmatrix}$$

- Create zeros under the newly created leading 1:

$$\begin{bmatrix}1 & -3 & 4 & -3 & 2 & 5\\0 & 1 & -2 & 2 & 1 & -3\\0 & 3 & -6 & 6 & 4 & -5\end{bmatrix} \xrightarrow{-3R_2+R_3} \begin{bmatrix}1 & -3 & 4 & -3 & 2 & 5\\0 & 1 & -2 & 2 & 1 & -3\\0 & 0 & 0 & 0 & 1 & 4\end{bmatrix}$$

- Now we have completed the top to bottom phase of the row reduction algorithm. In the next phase, we move from bottom to the top to create zeros above the leading 1s.

$$\begin{bmatrix}1 & -3 & 4 & -3 & 2 & 5\\0 & 1 & -2 & 2 & 1 & -3\\0 & 0 & 0 & 0 & 1 & 4\end{bmatrix} \xrightarrow{-R_3+R_2} \begin{bmatrix}1 & -3 & 4 & -3 & 2 & 5\\0 & 1 & -2 & 2 & 0 & -7\\0 & 0 & 0 & 0 & 1 & 4\end{bmatrix}$$

$$\begin{bmatrix}1 & -3 & 4 & -3 & 2 & 5\\0 & 1 & -2 & 2 & 0 & -7\\0 & 0 & 0 & 0 & 1 & 4\end{bmatrix} \xrightarrow{-2R_3+R_1} \begin{bmatrix}1 & -3 & 4 & -3 & \boxed{0} & -3\\0 & 1 & -2 & 2 & \boxed{0} & -7\\0 & 0 & 0 & 0 & 1 & 4\end{bmatrix}$$

*Create zeros above the leading 1 in row3*

# System of equations (CONT’D)

## Reduced Row Echelon Form

- Finally, create zeros above the leading 1 in the second row:

$$\begin{bmatrix} 1 & -3 & 4 & -3 & 0 & -3 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \xrightarrow{3R_2+R_1} \begin{bmatrix} 1 & \boxed{0} & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

- This completes the row reduction and the matrix is in RREF.

# System of equations (CONT'D)

## Reduced Row Echelon Form

- Example 2, consider the following linear system and its augmented matrix

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 8 \\ x_1 + 2x_2 + 4x_3 &= 8 \\ 3x_1 + 6x_2 + 9x_3 &= 12 \end{aligned} \qquad \begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 2 & 4 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix}$$

- After applying RREF, the resulting matrix is:

$$\begin{bmatrix} \boxed{1 \quad 2 \quad 3 \quad 4} \\ 0 \quad 0 \quad \boxed{1 \quad 4} \\ 0 \quad 0 \quad 0 \quad 0 \end{bmatrix}$$

$$x_1 + 2x_2 + 3x_3 = 4$$
$$\Rightarrow x_1 + 2x_2 = -8$$

$$x_3 = 4$$

- The RREF matrix contains 2 nonzero rows, but only 3 unknown variables. Therefore, this means that the solution set will contains $3 - 2 = 1$ ***free parameter***.
- From the above equation: $x_1 + 2x_2 = -8$ assume $\boxed{x_2 = t}$ the free parameter; $\Rightarrow \boxed{x_1 = -8 - 2t}$

# System of equations (CONT'D)

## Reduced Row Echelon Form

- Therefore, the solution set for the discussed linear system is:

$$x_1 = -8 - 2t$$
$$x_2 = t$$
$$x_3 = 4$$

- And as a rule, if a linear system has ***n*** unknown variables and the row reduced augmented matrix has ***r*** leading entries, then the number of free parameters d in the solution set is:

$$d = n - r$$

- Therefore, when performing back substitution, we have to set d of the unknown variables to arbitrary parameters.
- Because the number of leading entries r in the row reduced coefficient matrix determine the number of free parameters, we will refer to r as the rank of the coefficient matrix

# System of equations (CONT'D)

## Reduced Row Echelon Form

- The resulting RREF of the augmented matrix of a linear system will contain one of the distinct possibilities

1. The augmented matrix contains an inconsistent row (@side 27) → *The system is inconsistent and has no solution.*

2. All the rows of the augmented matrix are consistent and there are no free parameters → *The system is consistent and has a unique solution.*

3. All the rows of the augmented matrix are consistent and there are $d \geq 1$ variables that must be set to arbitrary parameters. → *The system is consistent and has infinitely many solutions. This happens when $r < n$ and thus $d \geq 1$ free parameters.*

- If we find the RREF form of the augmented matrix, ***the ones "1s" in the diagonal reflect the rank***.

# System of equations (CONT’D)

## Reduced Row Echelon Form

Non-singular matrix $\begin{bmatrix} 5 & 1 \\ 4 & -3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix}$

**Rank 2**

2 ones in the diagonal

Singular matrix $\begin{bmatrix} 5 & 1 \\ 10 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0.2 \\ 0 & 0 \end{bmatrix}$

**Rank 1**

1 one in the diagonal

Singular matrix $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

**Rank 0**

0 ones in the diagonal

# System of equations (CONT'D)
## Reduced Row Echelon Form / Row Echelon Form

| | | | | |
|---|---|---|---|---|
| 2 | * | * | * | * |
| 0 | 1 | * | * | * |
| 0 | 0 | 3 | * | * |
| 0 | 0 | 0 | -5 | * |
| 0 | 0 | 0 | 0 | 1 |

**Rank 5**

| | | | | |
|---|---|---|---|---|
| 3 | * | * | * | * |
| 0 | 0 | 1 | * | * |
| 0 | 0 | 0 | -4 | * |
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 |

**Rank 3**

- Zero rows at the bottom
- Each row has a pivot (leftmost non-zero entry)
- Every pivot is to the right of the pivots on the rows above
- Rank of the matrix is the number of pivots

With singular matrices we obtain inconsistent rows

# System of equations (CONT'D)

## Reduced Row Echelon Form / Row Echelon Form

**Matrix 1**

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 2 | 1 |
| 1 | 1 | 2 |

**Rank = 3**

**Matrix 2**

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 1 | 2 |
| 1 | 1 | 3 |

**Rank = 2**

**Matrix 3**

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 2 | 2 | 2 |
| 3 | 3 | 3 |

**Rank = 1**

**Matrix 4**

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 0 | 0 |
| 0 | 0 | 0 |

**Rank = 0**

**Row echelon forms**

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 0 | 1 | 0 |
| 0 | 0 | 1 |

**Number of pivots = 3**

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 0 | 0 | 1 |
| 0 | 0 | 0 |

**Number of pivots = 2**

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 0 | 0 | 0 |
| 0 | 0 | 0 |

**Number of pivots = 1**

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 0 | 0 |
| 0 | 0 | 0 |

**Number of pivots = 0**

# System of equations (CONT’D)

## Reduced Row Echelon Form / Row Echelon Form

| | | | | |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 1 |

**Rank 5**

| | | | | |
|---|---|---|---|---|
| 1 | * | 0 | 0 | * |
| 0 | 0 | 1 | 0 | * |
| 0 | 0 | 0 | 1 | * |
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 |

**Rank 3**

- Is in row echelon form
- Each pivot is a 1
- Any number above a pivot is 0
- Rank of the matrix is the number of pivots

# Matrix Multiplication (CONT'D)

## Vector-Vector Products (CONT'D)

- ***Inner-product*** or ***dot-product*** Multiplication:
  - Given two vectors $x, y \in \mathbb{R}^n$ the quantity $x^T y$, is sometimes called the inner product or dot product of the vectors, is a real number given by:

$$x^T y \in \mathbb{R} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ x_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i .$$

- ***Outer product***
  - Given vectors $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$ (vectors are not the same size), $xy^T \in \mathbb{R}^{m \times n}$, is called the outer product of the vector.

$$xy^T \in \mathbb{R}^{m \times n} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix} .$$

# Matrix Multiplication (CONT'D)

## Matrix-Vector Products (CONT'D)

- Given a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $x \in \mathbb{R}^{n}$, their product is a vector $y = Ax \in \mathbb{R}^{m}$
- There are two ways to look at matrix multiplication.
  - First: A is expressed by rows. And $y_i$ **is the inner product of the i-th row of A and x**.

$$y = Ax = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix} x = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{bmatrix}.$$

$x$ is a column vector

  - Alternatively, A can be written in column form, in this case, **y is a *linear combination* of the columns of A**, where the coefficients of the linear combination are given by the entries of x.

$$y = Ax = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \\ a_1 \\ \\ \end{bmatrix} x_1 + \begin{bmatrix} \\ a_2 \\ \\ \end{bmatrix} x_2 + \ldots + \begin{bmatrix} \\ a_n \\ \\ \end{bmatrix} x_n$$

# Matrix Multiplication (CONT'D)

## Matrix-Vector Products (CONT'D)

- So far, we have been multiplying on the right by a column vector, but it is also possible to multiply on the left by a row vector. This can be written as $y^T = x^T A$ for $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$
    - A matrix can be treated as a group of column vectors.

$$y^T = x^T A = x^T \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} x^T a_1 & x^T a_2 & \cdots & x^T a_n \end{bmatrix}$$

- $\Rightarrow$ the i<sup>th</sup> entry of $y^T$ **is equal to the inner product of $x$ and the i<sup>th</sup> column of $A$**.
- By expressing $A$ as rows, $y^T$ **is a linear combination of the row vectors of $A$** where the coefficients are stored in $x$.

$$\begin{aligned} y^T &= x^T A \\ &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix} \\ &= x_1 \begin{bmatrix} — & a_1^T & — \end{bmatrix} + x_2 \begin{bmatrix} — & a_2^T & — \end{bmatrix} + \ldots + x_n \begin{bmatrix} — & a_n^T & — \end{bmatrix} \end{aligned}$$

# Matrix Multiplication (CONT'D)

## Matrix-Matrix Products (CONT'D)

- Matrix multiplication can be seen in 4 different ways
  - **Can be seen as a set of vector-vector products.**
  - If we represent A by rows and B by columns:

  The (i, j)$^{th}$ entry of C is equal to the inner product of the i$^{th}$ row of A and the j$^{th}$ col of B. Symbolically, this looks like the following:

$$C = AB = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} a_1^T b_1 & a_1^T b_2 & \cdots & a_1^T b_p \\ a_2^T b_1 & a_2^T b_2 & \cdots & a_2^T b_p \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T b_1 & a_m^T b_2 & \cdots & a_m^T b_p \end{bmatrix}$$

$\mathbb{R}^{m\times n}$ $\quad$ $\mathbb{R}^{n\times p}$

$a_i \in \mathbb{R}^n$ $\quad$ $b_j \in \mathbb{R}^n$ $\quad$ $\Rightarrow$ All inner products make sense

# Matrix Multiplication (CONT'D)

## Matrix-Matrix Products (CONT'D)

- Matrix multiplication can be seen in 4 different ways
  - **Can be seen as a set of vector-vector products**
  - On the other hand, if we represent **A by columns, and B by rows**

    ⇒ This is equivalent to the sum of outer products of the vectors $a_i$ and $b_i$.

$$C = AB = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} — & b_1^T & — \\ — & b_2^T & — \\ & \vdots & \\ — & b_n^T & — \end{bmatrix} = \sum_{i=1}^{n} a_i b_i^T$$

$\mathbb{R}^{m \times n}$ $\quad$ $\mathbb{R}^{n \times p}$

$a_i \in \mathbb{R}^m$ $\quad$ $b_i \in \mathbb{R}^p$ $\quad \Rightarrow$ The dimension of the outer product Is m x p.

# Matrix Multiplication (CONT'D)

## Matrix-Matrix Products (CONT'D)

- Finally, matrix-matrix multiplication can be seen as **a set of matrix-vector products between A and the columns of B**.

$$C = AB = A \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ Ab_1 & Ab_2 & \cdots & Ab_p \\ | & | & & | \end{bmatrix}$$

- Alternatively, we can view A by rows and B by columns.

$$C = AB = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix} B = \begin{bmatrix} — & a_1^T B & — \\ — & a_2^T B & — \\ & \vdots & \\ — & a_m^T B & — \end{bmatrix}$$

# Matrix Multiplication (CONT'D)
## Basic Properties

- Matrix multiplication is associative: $(AB)C = A(BC)$
- Matrix multiplication is distributive: $A(B + C) = AB + AC$
- Matrix multiplication is, in general, not commutative; that is, it can be the case that $AB \neq BA$.

# LU Factorization

- The matrix form of one elimination step, assume we want to apply
  - $2 * eq1 \; - eq2$:

$$E = \begin{bmatrix} 1 & 0 & 0 \\ -\mathbf{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ has } Ib = b$$

$$Eb = \begin{bmatrix} 1 & 0 & 0 \\ -\mathbf{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} = \begin{bmatrix} 5 \\ -\mathbf{12} \\ 9 \end{bmatrix}.$$

  - The elementary matrix $E_{ij}$ subtracts $l$ times row $j$ from row $i$.
    - E.g., subtract 3 times equation 1 from equation 3.

$$E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\ell & 0 & 1 \end{bmatrix}$$

# LU Factorization

- Assume matrix F represent the second forward elimination step. E.g., $F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mathbf{1} & 0 & 1 \end{bmatrix}$
- Now, both steps can be applied at once by multiple E and F. $EF = \begin{bmatrix} 1 & 0 & 0 \\ -\mathbf{2} & 1 & 0 \\ \mathbf{1} & 0 & 1 \end{bmatrix}$
- Assume needing to apply another elimination step, this could be represented by G. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \mathbf{1} & 1 \end{bmatrix}$
- Apply those 3 elimination steps take A to U "upper triangular matrix"

$$\textbf{From } A \textbf{ to } U \qquad GFE = \begin{bmatrix} 1 & & \\ & 1 & \\ & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ 1 & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ -2 & 1 & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -2 & 1 & \\ -1 & 1 & 1 \end{bmatrix}$$

# LU Factorization

- The inverse of each elementary step is easy, the inverse of subtraction is addition

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E^{-1}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$GFE = \begin{bmatrix} 1 & & \\ & 1 & \\ & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ 1 & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ -2 & 1 & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -2 & 1 & \\ -1 & 1 & 1 \end{bmatrix}$$

$$E^{-1}F^{-1}G^{-1} = \begin{bmatrix} 1 & & \\ \mathbf{2} & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ \mathbf{-1} & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & \mathbf{-1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ \mathbf{2} & 1 & \\ \mathbf{-1} & \mathbf{-1} & 1 \end{bmatrix} = L$$

$$GFEA = U \Rightarrow A = E^{-1}F^{-1}G^{-1}U = LU$$

$$\boldsymbol{\Rightarrow A = E^{-1}F^{-1}G^{-1}U = LU}$$

$$Ux = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix} = c.$$

# LU Factorization

- $LU$ factorization allows us to solve a system of linear equations in more efficient way.
- First, apply Gaussian elimination to find $A = LU$.
- Instead of solving $Ax = b$,
  - Replace $A$ by $LU$ to get: $LUx = b$
  - Let $y = Ux$,
    - Now, both $y$ and $x$ are unknowns. $y = [y_1, y_2, \dots, y_n]$
  - But $L$ and $b$ are known $\Rightarrow$ solve $y$ in $Ly = b$ using forward substitution.
  - After solving $y$.
  - Then solve for $x$ in $Ux = y$ using back substitution.
- This way, we factorize $A$ to $LU$ once, and then apply forward and backward substitution for any $x$.
  - “This is why its efficient”

# LU Factorization

- Other benefits of $LU$ factorization include:
  - **A fast way for computing the determinants.**
  - The determinant of a triangular matrix is the product of the elements across the main diagonal.
  - Since $A = LU$, then $det(A) = \det(LU) = \det(L) * \det(U)$
- Directly calculating the inverse of a matrix can be computationally intensive, especially for large matrices.
- LU factorization simplifies the process of finding the inverse of a matrix. Here's how it can be beneficial:
  - Suppose we have a matrix $A$ and we want to find its inverse $A^{-1}$.
  - If we perform $LU$ factorization on $A$, we have $A = LU$.
  - To find $A^{-1}$, we need to solve $AA^{-1} = I$ for $A^{-1}$, where $I$ is the identity matrix.
  - Substituting $A = LU$ into the equation, we get $LUA^{-1} = I$.
  - Let's denote $A^{-1} = [x_1, x_2, \dots, x_n]$ where $x_i$ are the columns of $A^{-1}$.
  - The equation then becomes $LU * x_i = e_i$, where $e_i$ are the columns of $I$.

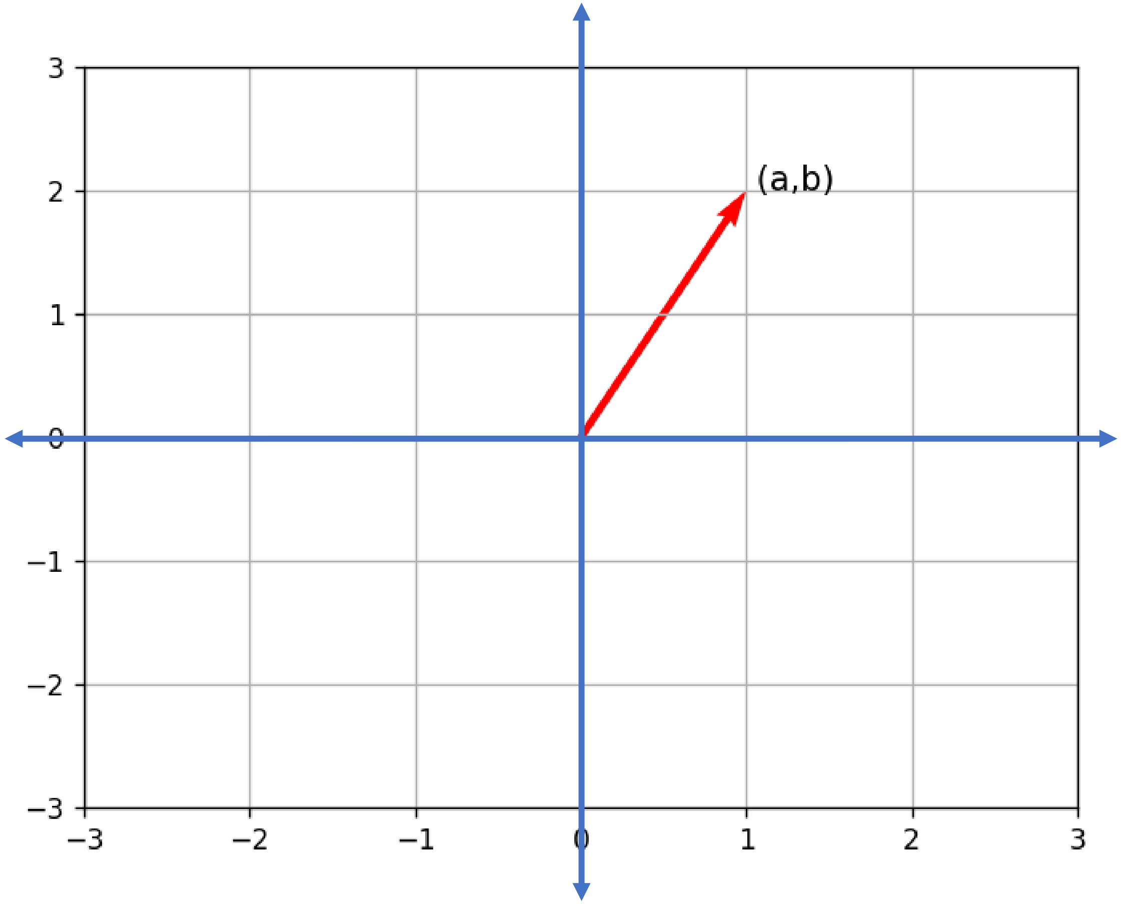
# LU Factorization

- Other benefits of $LU$ factorization include:
- The computation of $A^{-1}$ has been transformed into solving $n$ systems of linear equations $LU * x_i = e_i$, for $i = 1, \ldots n$. This is more efficient because:
  - Solving a system of equations using LU factorization is faster compared to direct inversion, especially for large matrices.
  - L and U are triangular matrices, and solving a system of linear equations is straightforward when the coefficient matrix is triangular (forward and backward substitution)

# Vector and Linear Transformation

## Norms

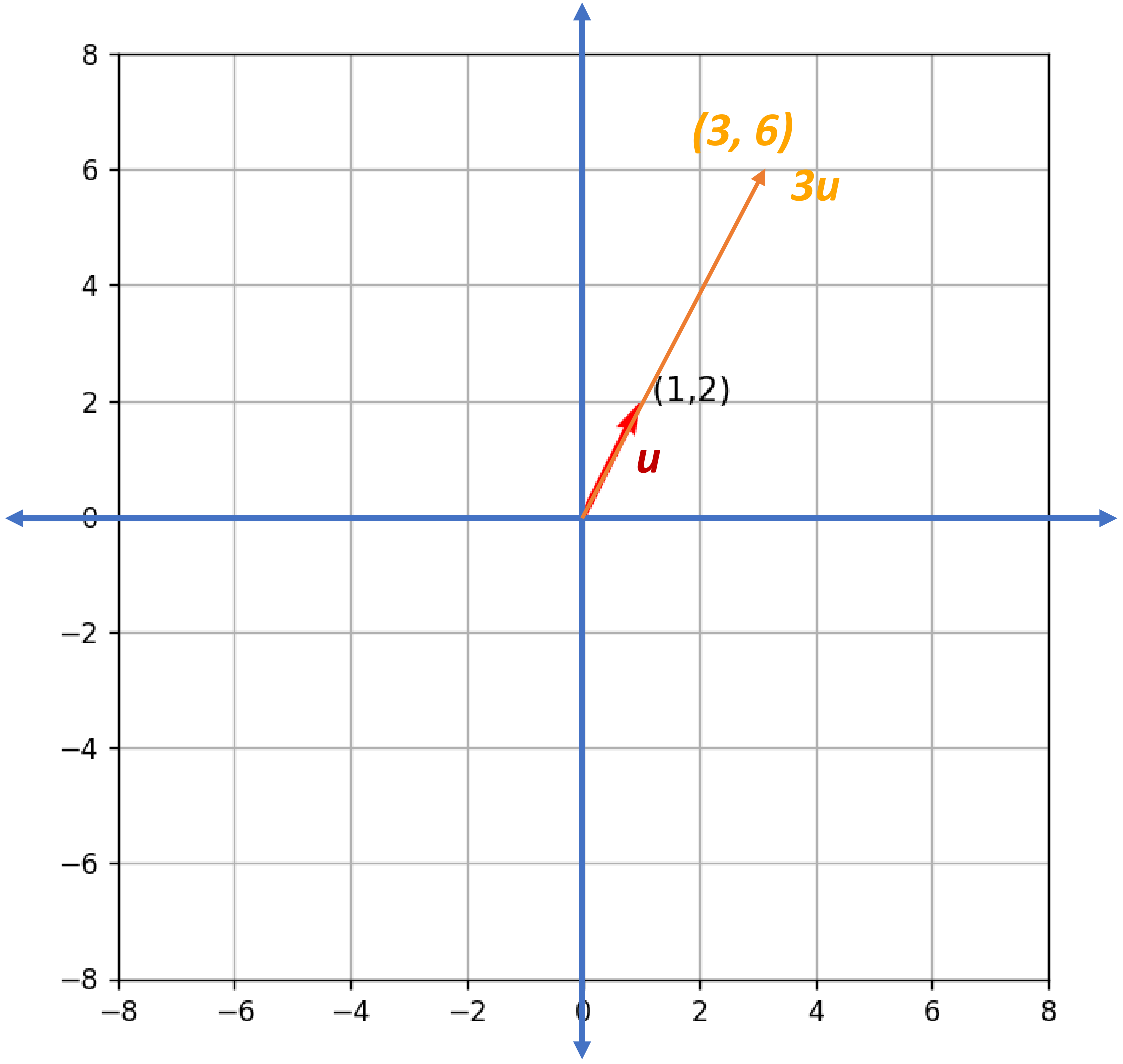
- $L_2 norm = |(a,b)|_2 = \sqrt{a^2 + b^2}$
- $L_1 norm = |(a,b)|_1 = |a| + |b|$



# Vector and Linear Transformation

## Norms

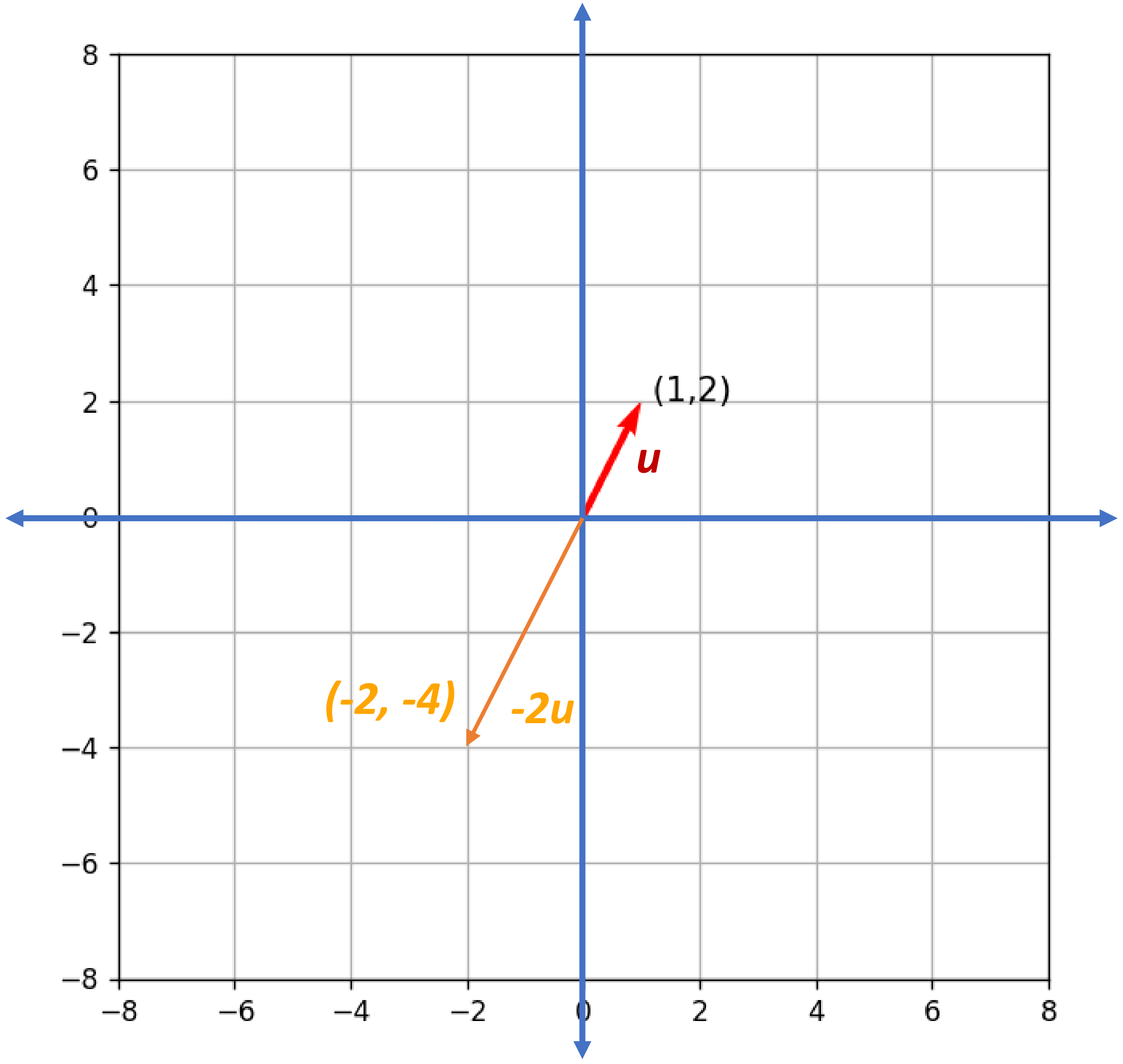
- $u = (1,2)$
- $\lambda = 3$
- $\lambda u = (3,6)$



# Vector and Linear Transformation

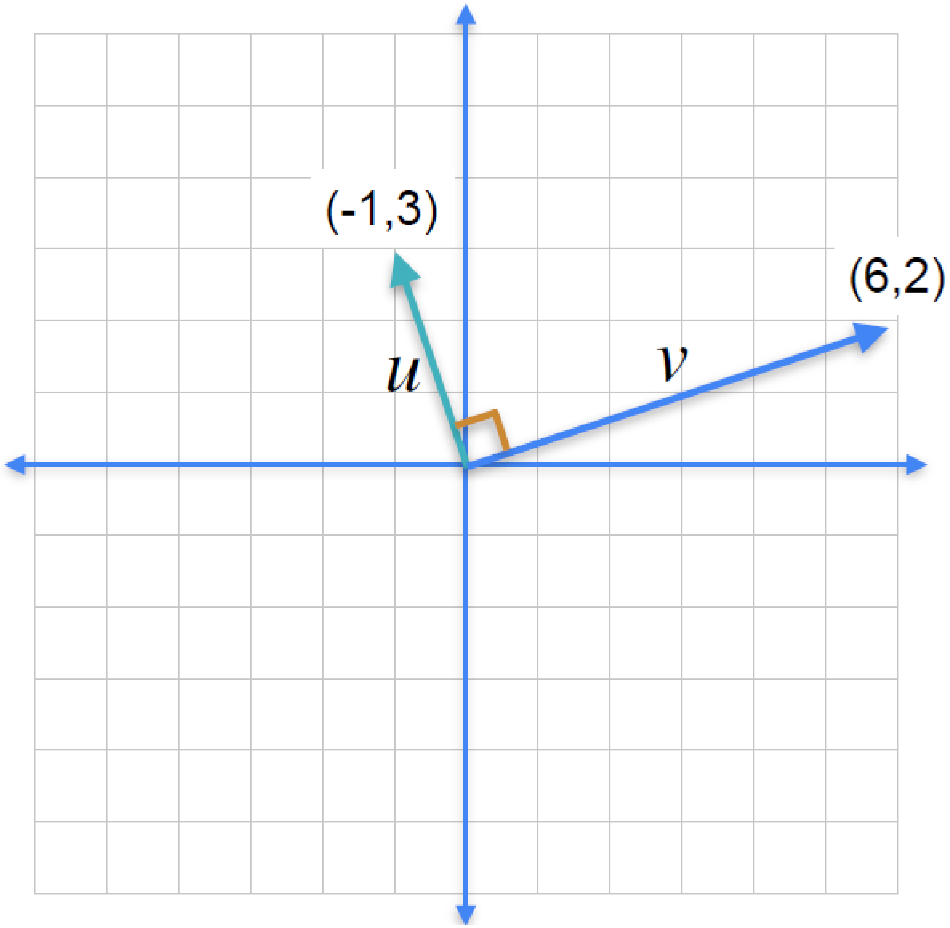
## Norms

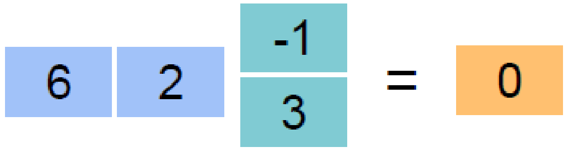
- $u = (1,2)$
- $\lambda = -2$
- $\lambda u = (-2, -4)$



# Vector and Linear Transformation

## Geometric Dot Product

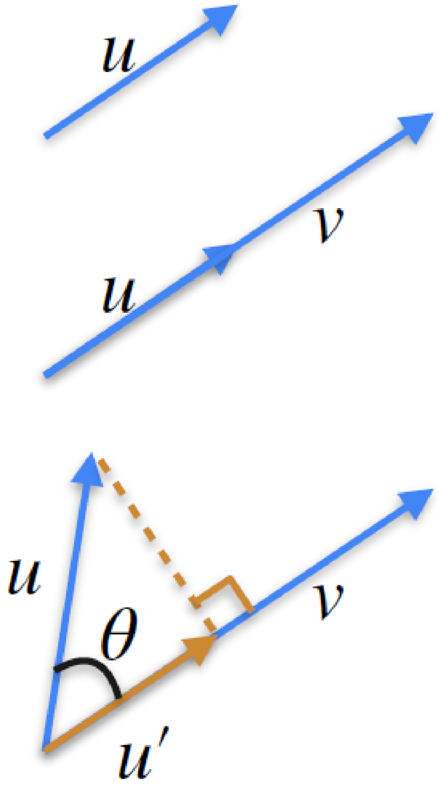



$$\langle u, v \rangle = 0$$

# Vector and Linear Transformation

## Geometric Dot Product



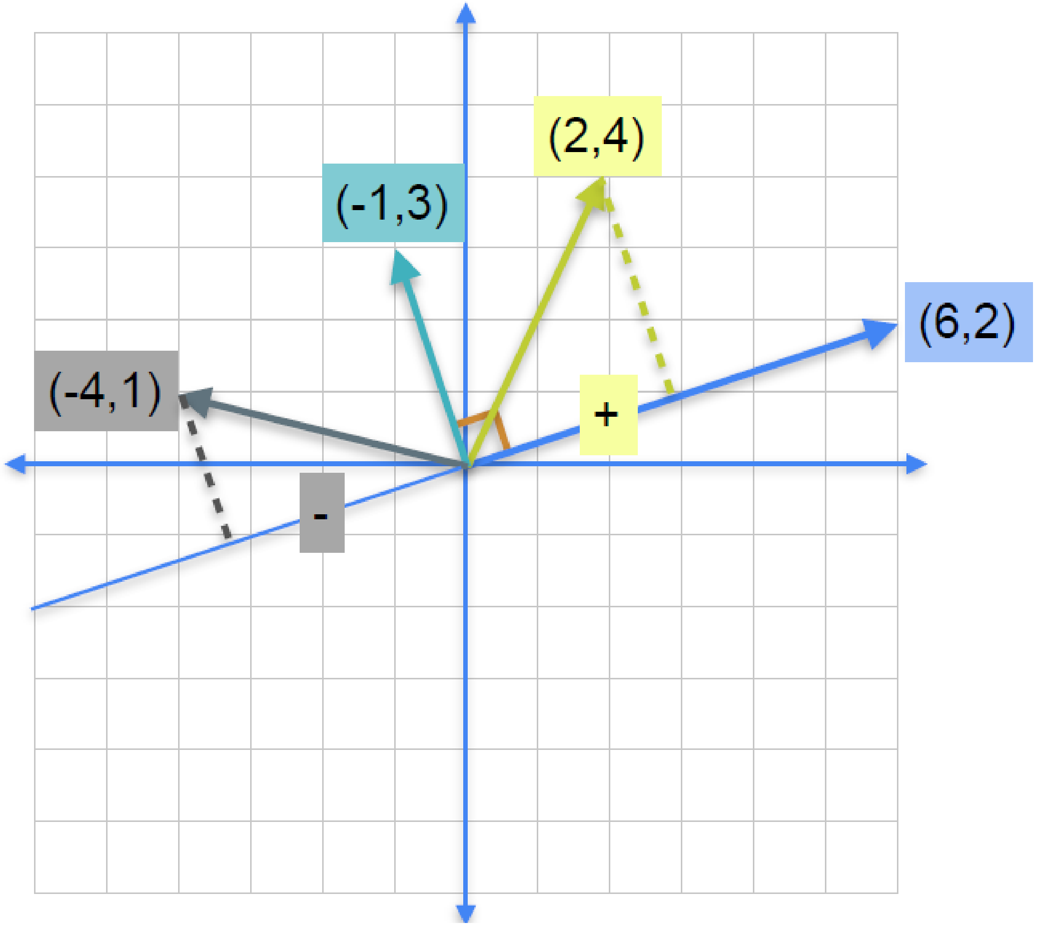

$$\langle u, u \rangle = |u|^2 = |u| \cdot |u|$$

$$\langle u, v \rangle = |u| \cdot |v|$$

$$\begin{aligned}\langle u, v \rangle &= |u'| \cdot |v| \\ &= |u||v|\cos(\theta)\end{aligned}$$

# Vector and Linear Transformation

## Geometric Dot Product



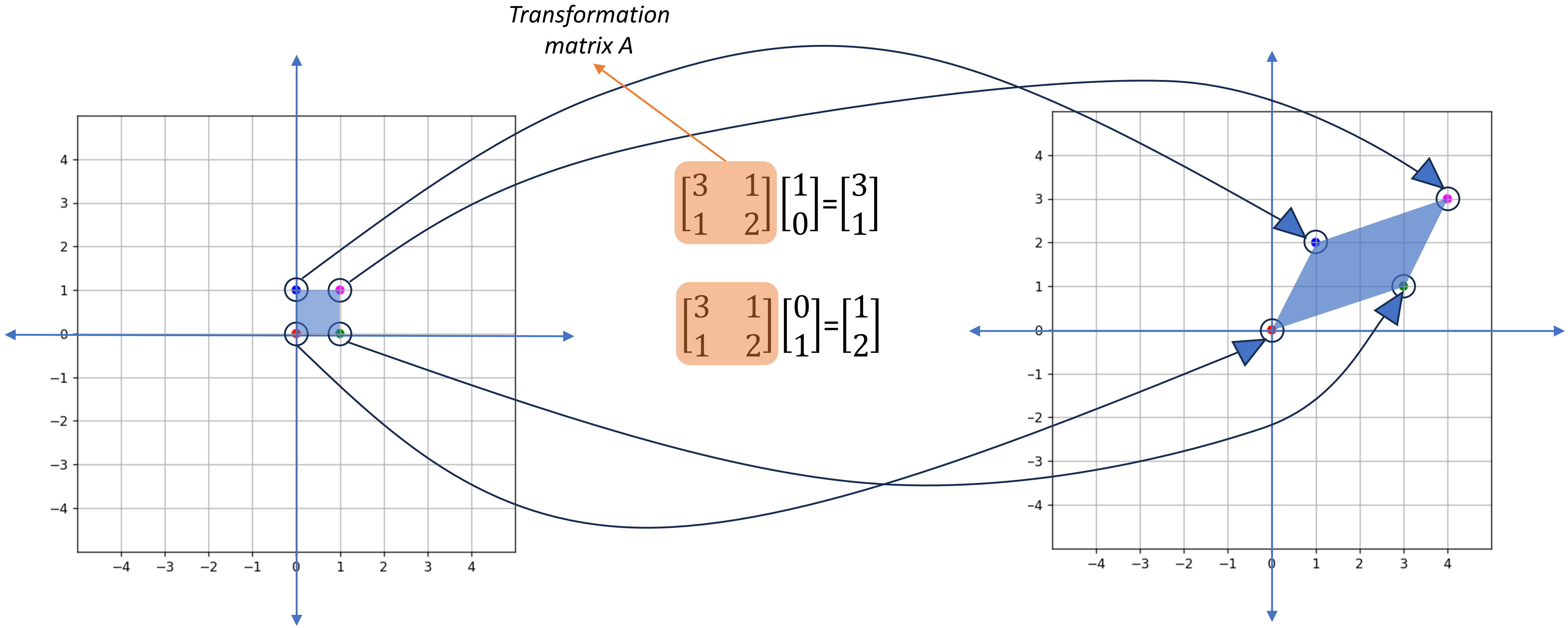

$$\begin{bmatrix} 6 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 20$$ Positive

$$\begin{bmatrix} 6 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 6 & 2 \end{bmatrix} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = -22$$ Negative

# Vector and Linear Transformation

## Matrices as Linear Transformation

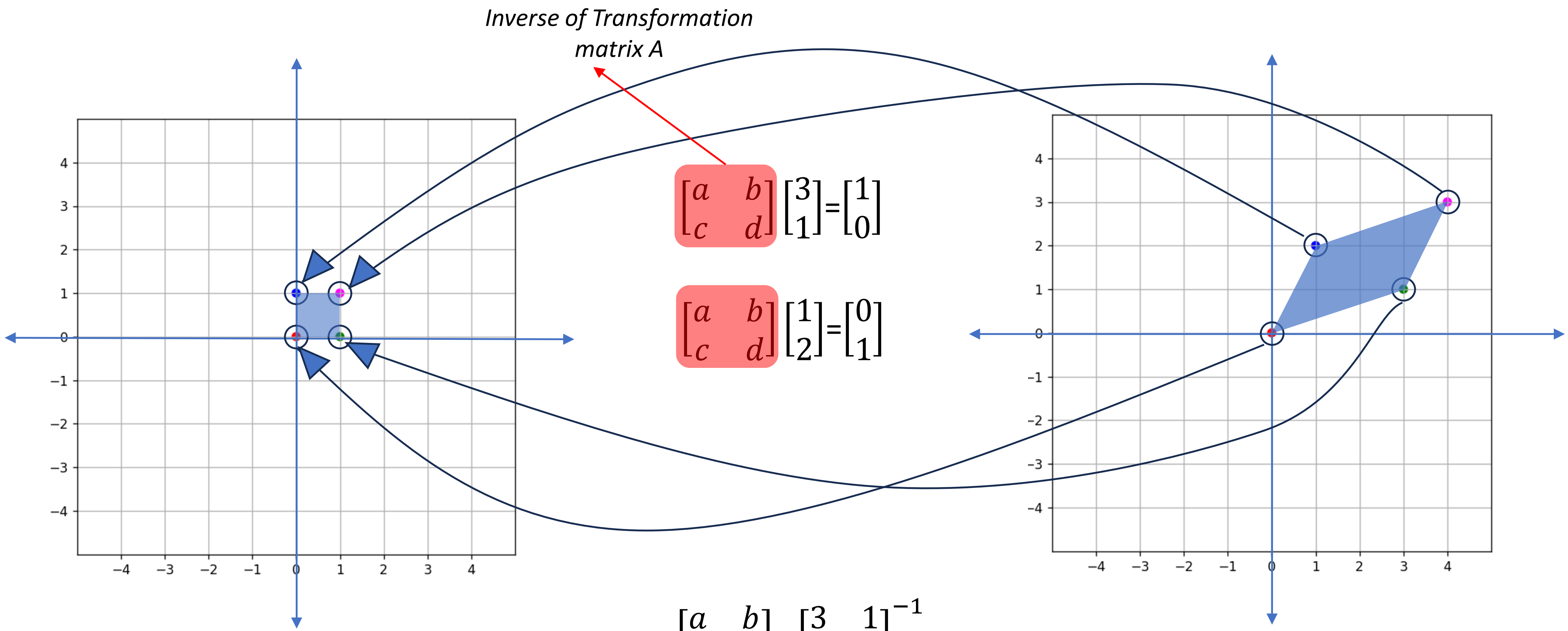



$$\begin{bmatrix}3 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}$$

$$\begin{bmatrix}3 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\2\end{bmatrix}$$

# Vector and Linear Transformation

## Matrices as Linear Transformation




$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix}=\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix}=\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

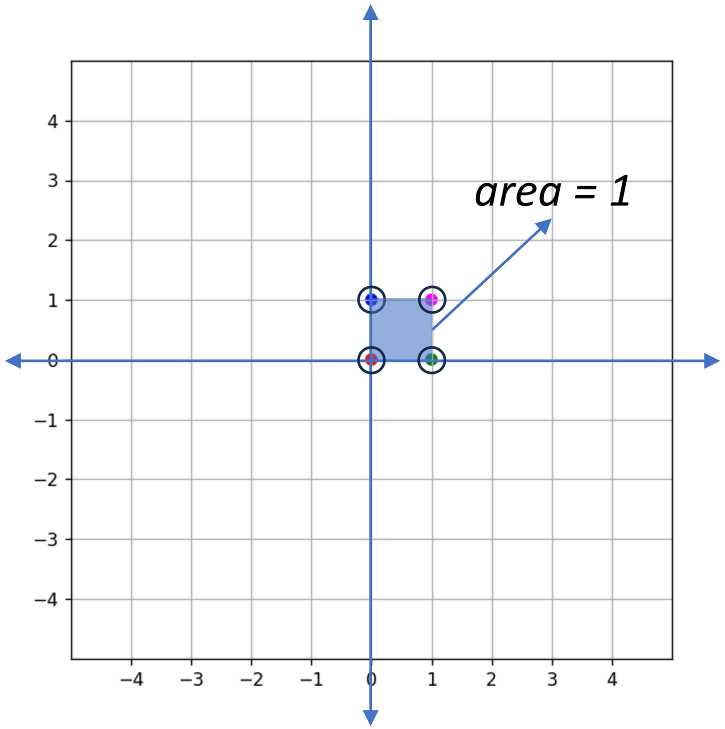
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}=\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$$

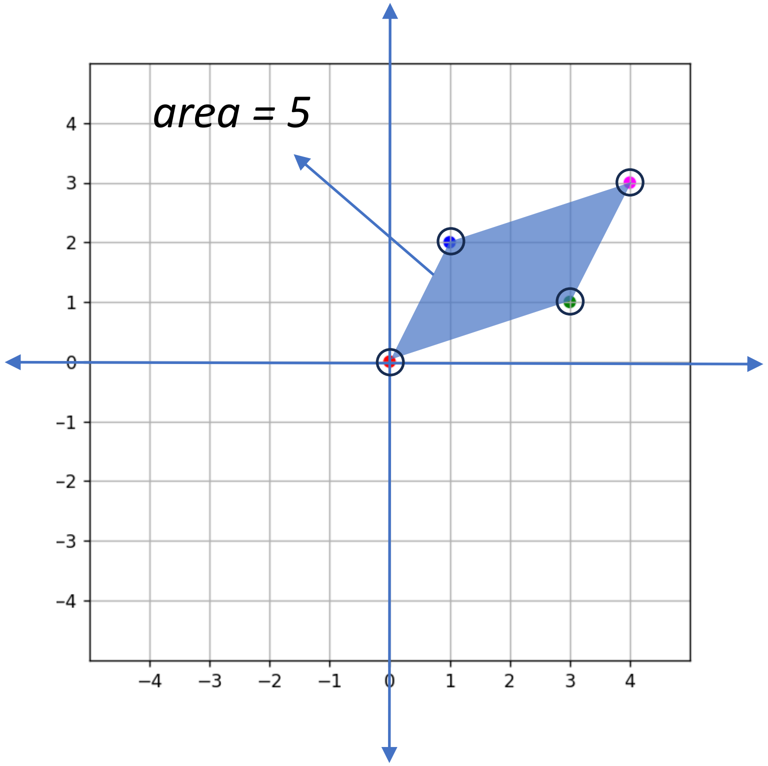
# Vector and Linear Transformation

## Determinant as an Area

- In the context of linear transformation, the determinant can be seen as being the area or as a volume of the image of the basis

$$\det\left(\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}\right) = 3.2 - 1.1 = 5$$



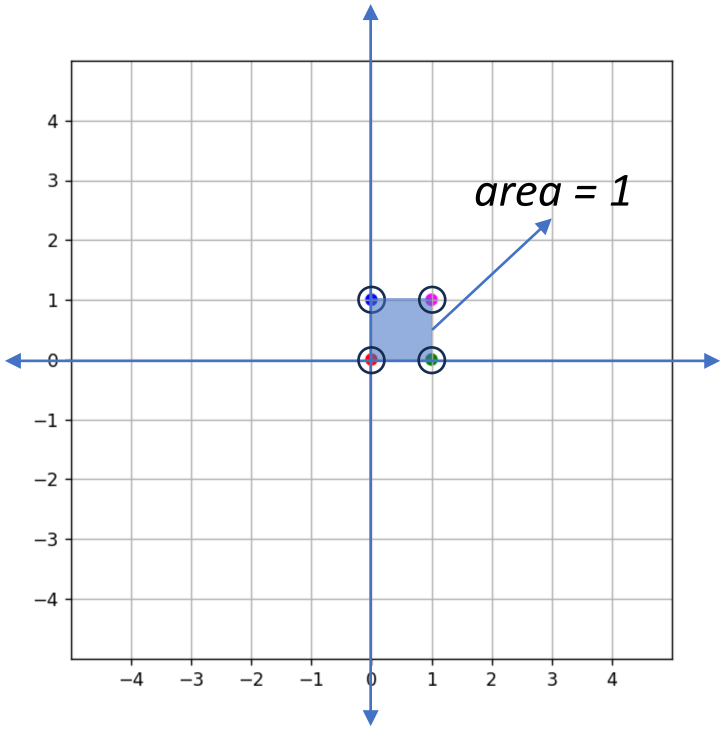



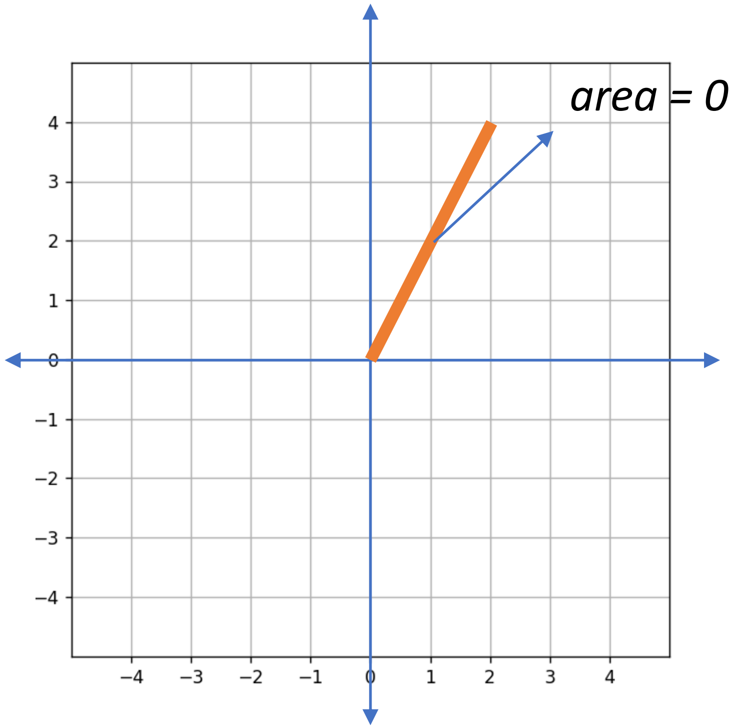
# Vector and Linear Transformation

## Determinant as an Area

- In the case of singularity, this area is equal to zero (det = 0)

$$\det\left(\begin{bmatrix}1 & 1\\2 & 2\end{bmatrix}\right) = 1.2 - 1.1 = 0$$

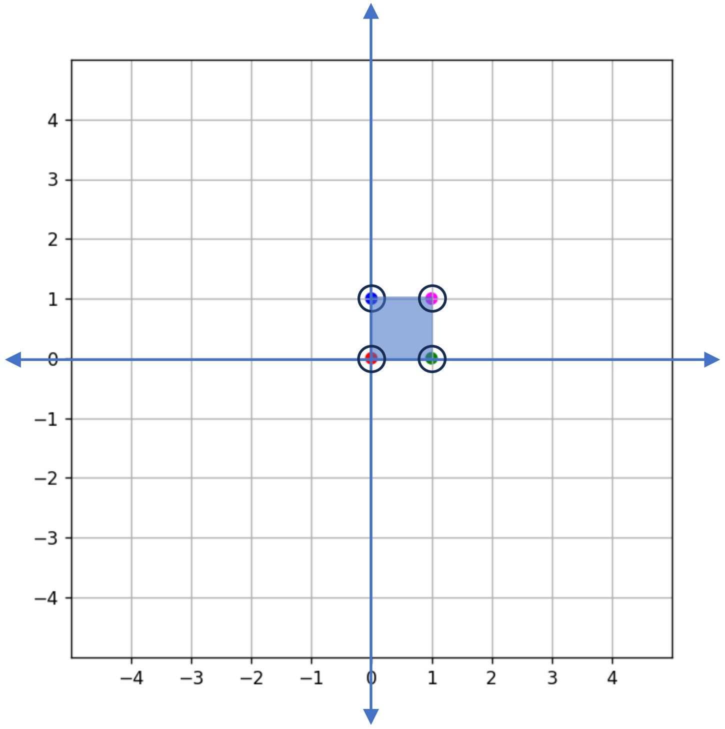


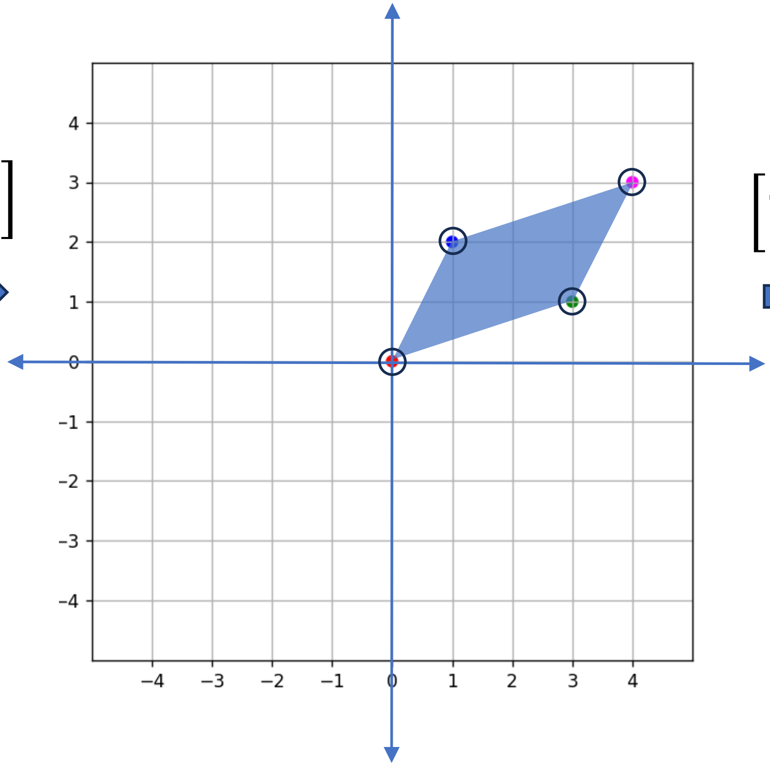
# Identity and Inverse Matrices

- Linear algebra offers a powerful tool called matrix inversion that allows us to analytically solve $Ax = b$ for any $A$
- An identity matrix is a matrix that does not change any vector when we multiply that vector by that matrix.
- How to solve for the inverse of matrix $A$?
- If $A$ is a transformation matrix that takes $x$ to $b$: $Ax = b \quad \Rightarrow \quad x = A^{-1}b$ ; $A^{-1}$ takes $b$ back to $x$.
- Rule: $AA^{-1} = I$

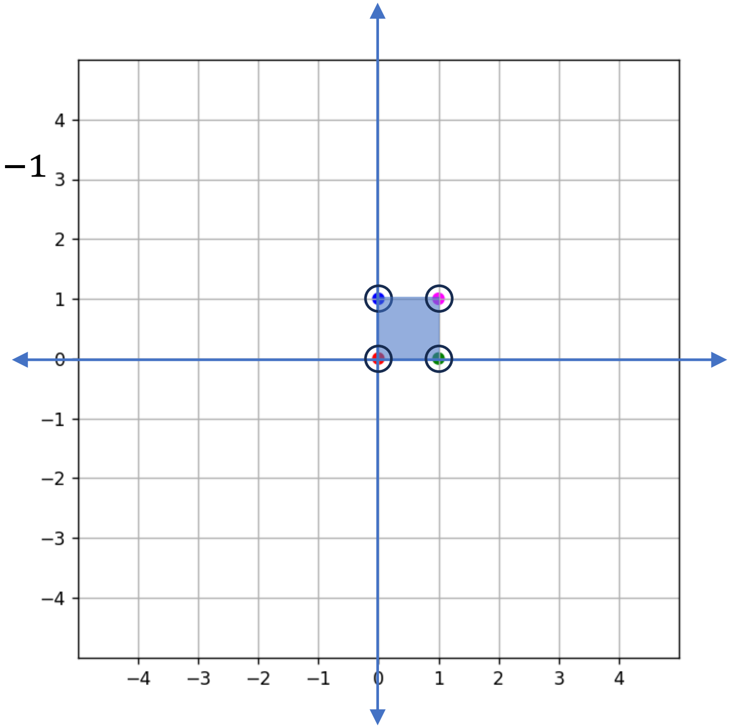
# Identity and Inverse Matrices



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



$$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$$



$$\mathrm{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

# Identity and Inverse Matrices

- But which matrices has inverse?
  - The matrix should be ***invertible/Non-singular*** $\Rightarrow$ There should exist unique solution $\Rightarrow$ for every $b$ in $Ax = b$ there exist one $x$. Also,
  - Note that not all matrices have inverses. Non-square matrices, for example, do not have an inverse.
- How to find the inverse?
  - Solve the system of linear equation resulting from $AA^{-1} = I$.

$$\Rightarrow \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} . \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This comes down to solving:

$$\begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} a \\ c \end{bmatrix} = 1 \qquad 3a + 1c = 1$$

$$\begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} b \\ d \end{bmatrix} = 0 \qquad 3b + 1d = 0$$

$$\Rightarrow$$

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ c \end{bmatrix} = 0 \qquad 1a + 2c = 0$$

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} b \\ d \end{bmatrix} = 1 \qquad 1a + 2c = 1$$

$\Rightarrow$ We now have a system of 4 linear equations of 4 unknown.

# Special Kind of Matrices and Vectors

- ***Diagonal matrices*** consist mostly of zeros and have non-zero entries only along the main diagonal.
- To represent a square matrix, use $diag(v)$.
- Diagonal matrices are of interest in part because multiplying by a diagonal matrix is very computationally efficient.
  - $diag(v)x$ scale each element $x_i$ by $v_i$.
- Inverting a square diagonal matrix is also efficient (only if every diagonal entry is nonzero).
  - $diag(v)^{-1} = diag(\left[\frac{1}{v_1}, \dots, \frac{1}{v_n}\right]^T)$
- It is also possible to construct a rectangular diagonal matrix (but those do not have an inverse).
- A ***symmetric matrix*** is a matrix that is equal to its own transpose.

$$A = A^T$$

# Special Kind of Matrices and Vectors

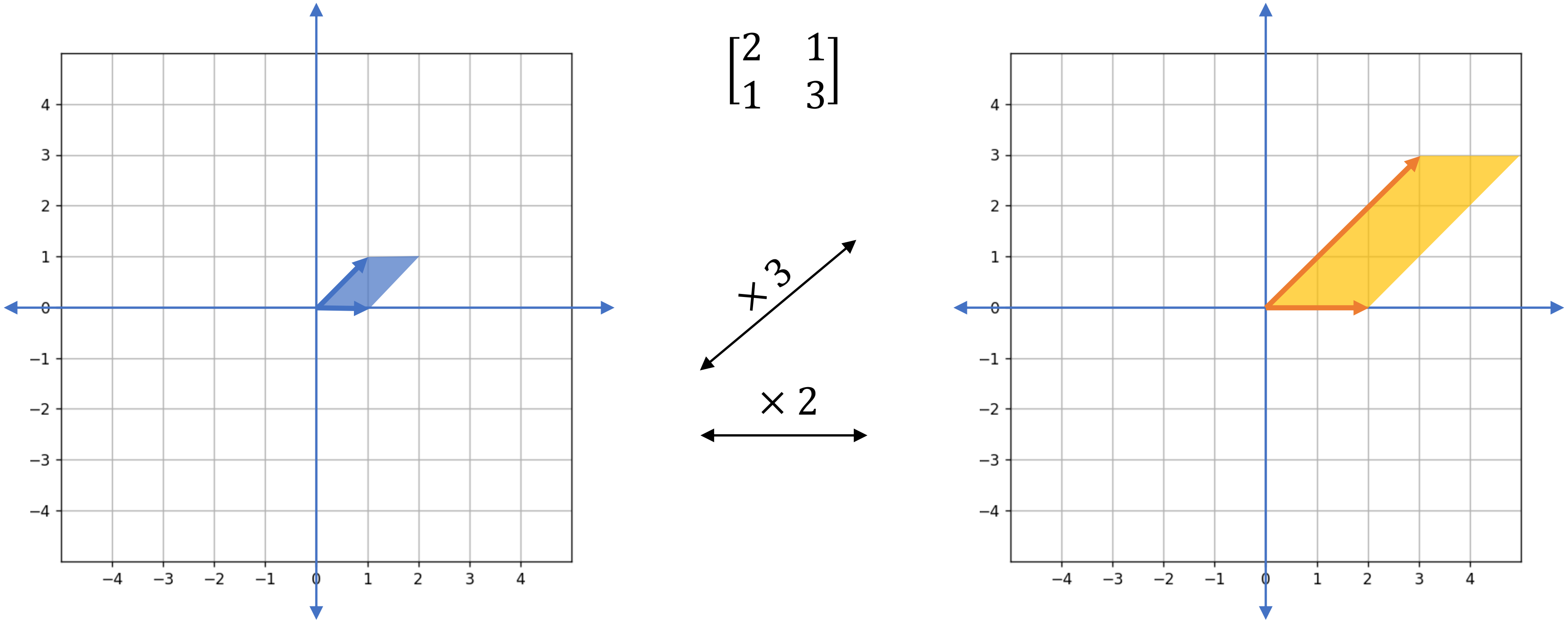
- A vector $x$ and a vector $y$ are orthogonal to each other if $x^T y = 0$ where both are nonzero.
- In $R^n$ at most $n$ vectors may be mutually orthogonal with nonzero norm.
  - When orthogonal vectors have unit norm, we call them orthonormal.
- An orthogonal matrix is a square matrix whose rows are mutually orthonormal and whose columns are mutually orthonormal.
  - $A^T A = AA^T = I \quad \Rightarrow A^T = A^{-1}$
  - Therefore, orthogonal matrices are of interest because their inverse is very cheap to compute.

# Eigendecomposition

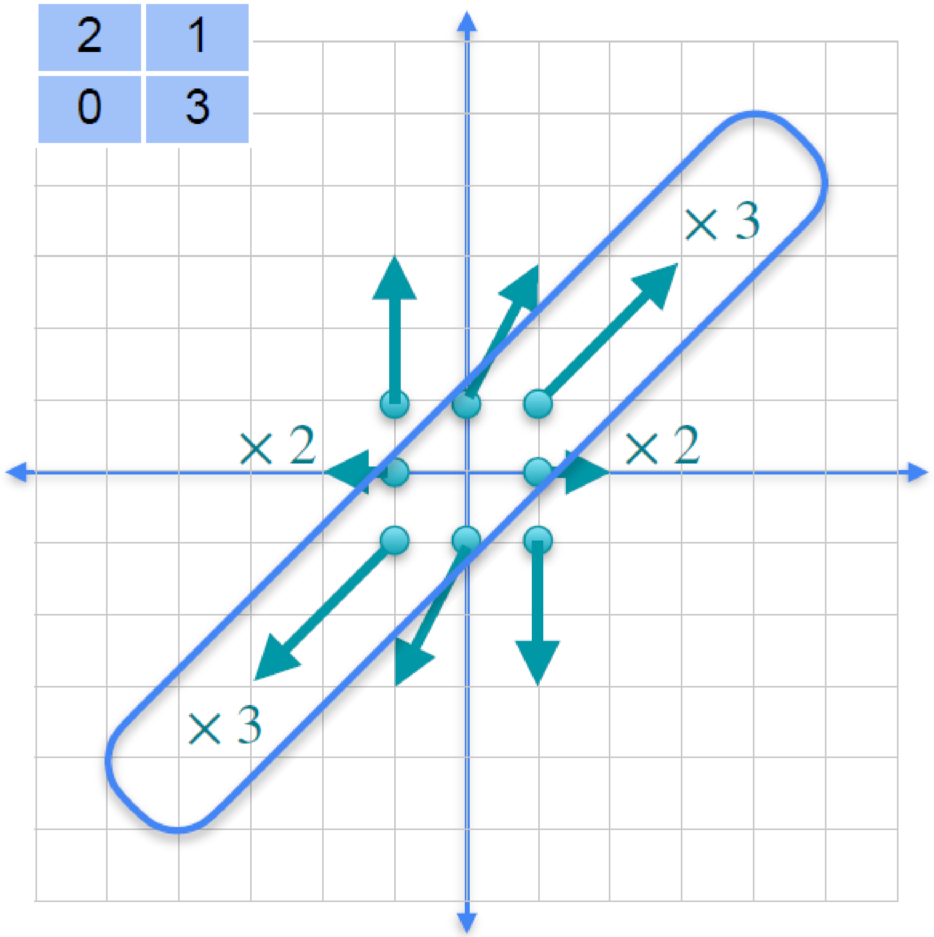
- Similar to how numbers can be decomposed into prime factors, we can decompose matrices in a way that show us information about their functional properties.
- One of the most widely used kinds of matrix decomposition is called ***eigndecomposition***.
    - It decomposes a matrix into a set of *eigenvalues* and *eigenvectors*.
- The eigenvector of a square matrix $\boldsymbol{A}$ is a nonzero vector $\boldsymbol{v}$ such that when $\boldsymbol{A}$ is multiplied by $\boldsymbol{v}$, $\boldsymbol{A}$ will only alter the scale of $\boldsymbol{v}$.
    - $\boldsymbol{Av} = \boldsymbol{\lambda v}$
    - $\boldsymbol{\lambda}$ is known as the eigenvalue corresponding to the eigenvector $\boldsymbol{v}$.
    - If $\boldsymbol{v}$ is an eigenvector, so is $\boldsymbol{sv}$ for $\boldsymbol{s} \in \mathbb{R}$. Yet, $\boldsymbol{sv}$ still has the same eigenvalues.
    - Therefore, we only consider unit eigenvectors.

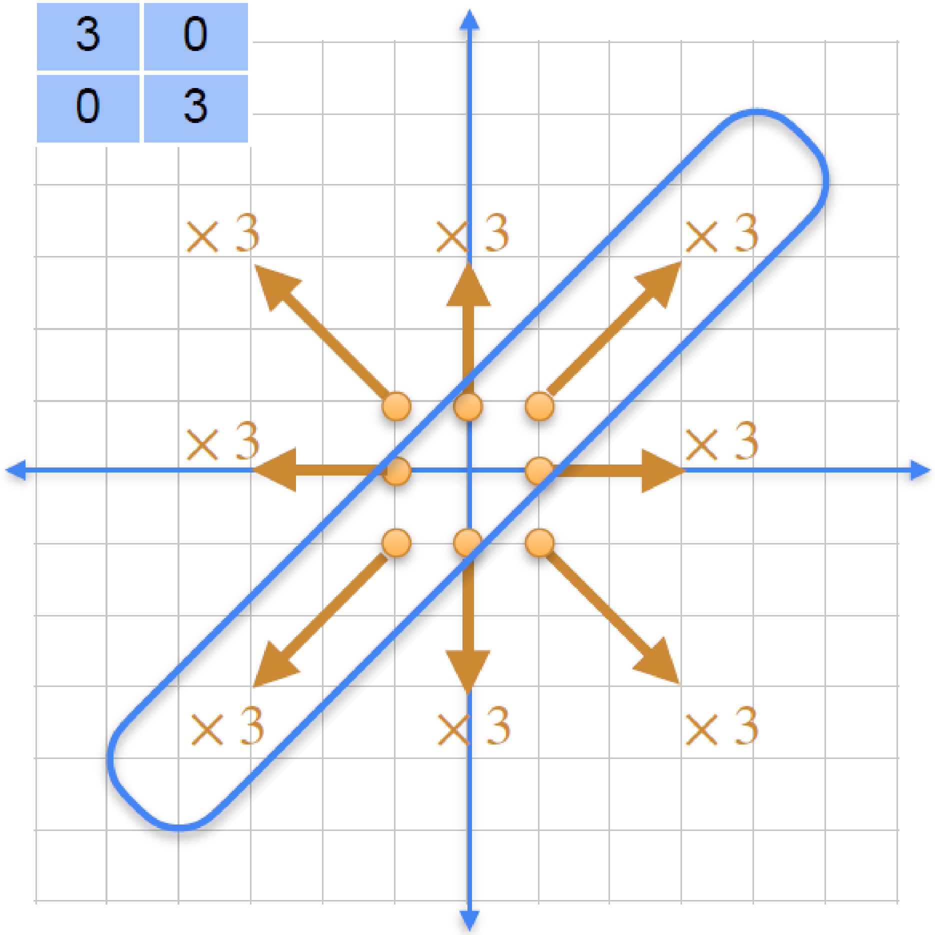
# Eigendecomposition

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

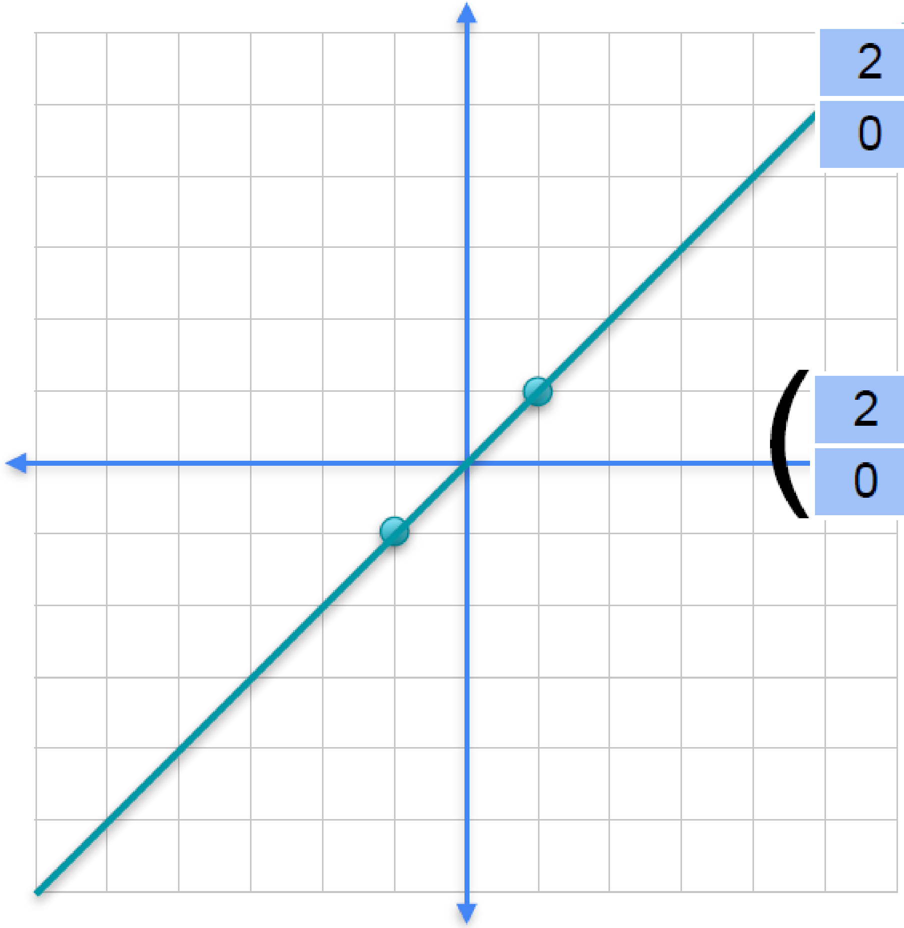


# Eigendecomposition



# Eigendecomposition



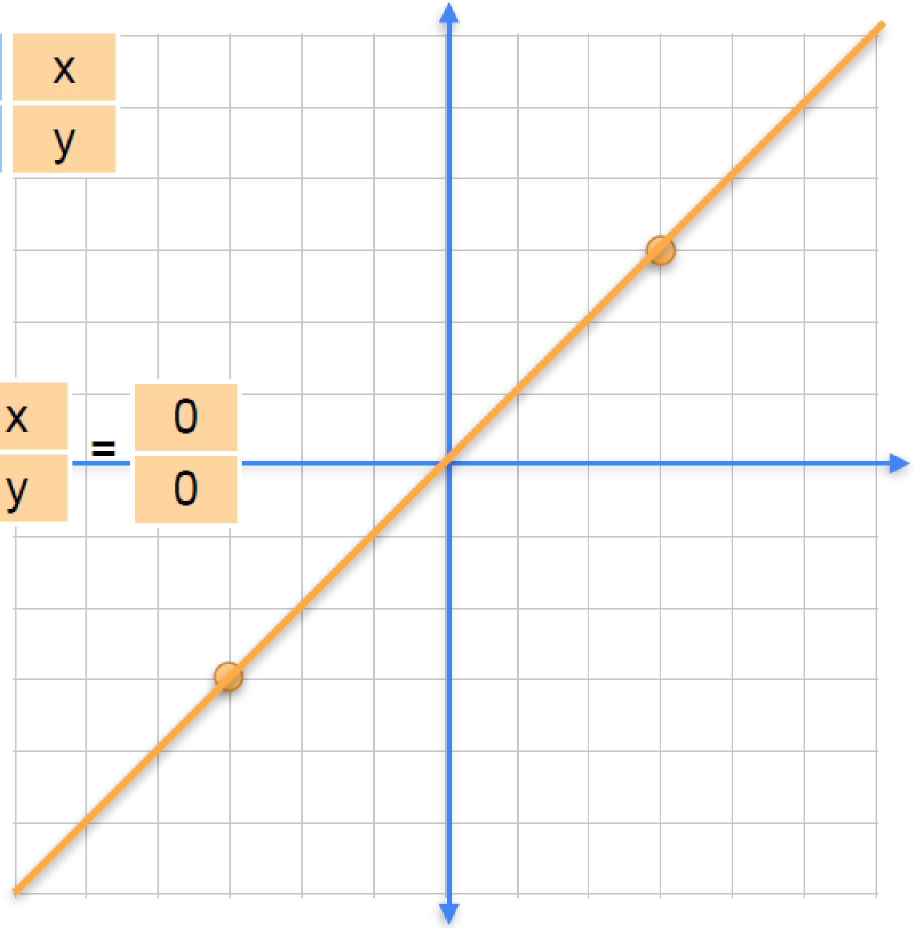
$$\begin{bmatrix}2&1\\0&3\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}3&0\\0&3\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

For infinitely many points

$$\left(\begin{bmatrix}2&1\\0&3\end{bmatrix} - \begin{bmatrix}3&0\\0&3\end{bmatrix}\right)\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

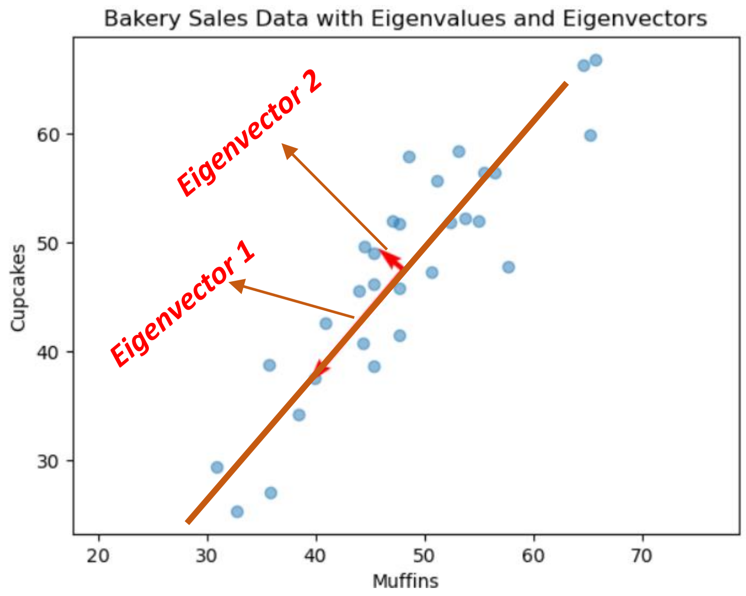
$$\begin{bmatrix}-1&1\\0&0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

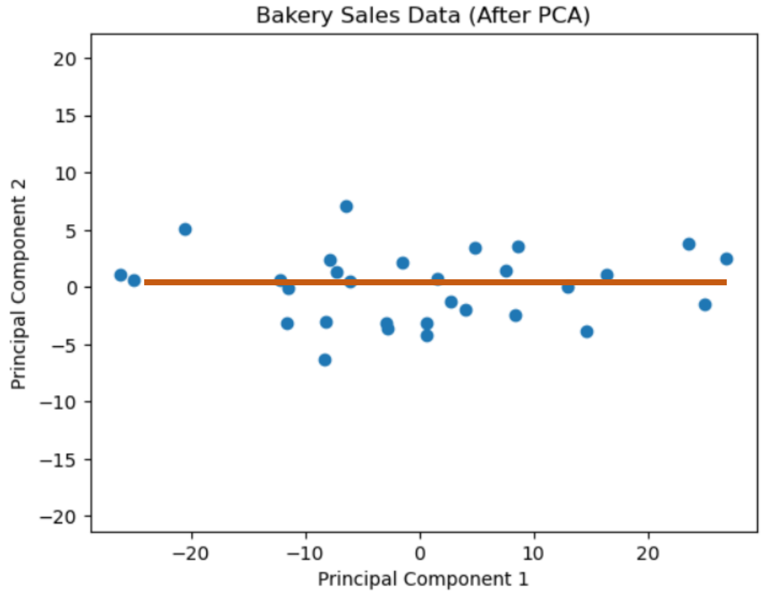
$$\det\begin{bmatrix}-1&1\\0&0\end{bmatrix} = 0$$

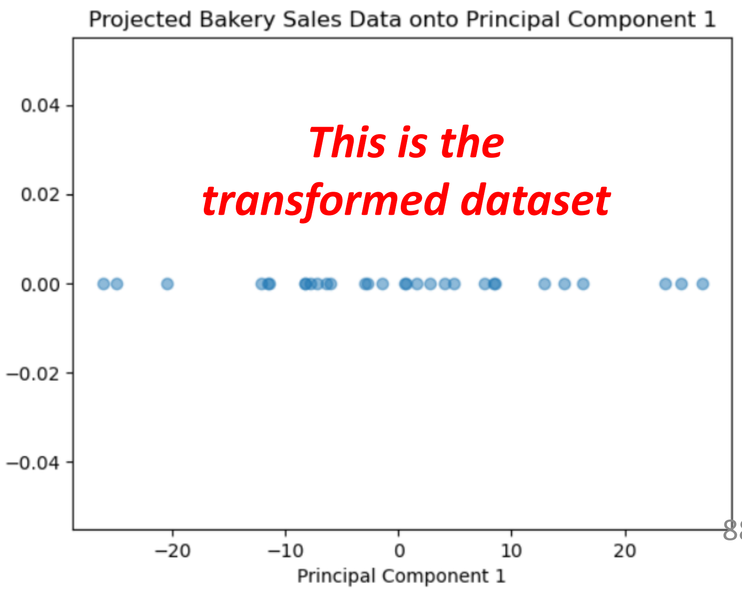


# Principle Component Analysis

- Imagine we have the following data points representing sales of muffins and cupcakes
- Here is an example on what the PCA will do.
- To simplify the dataset, we can look at the same dataset from the point of view of the line that fits them the most.
- And this is what PCA do: it will find the line and then take a look at the dataset from the point of view of the line



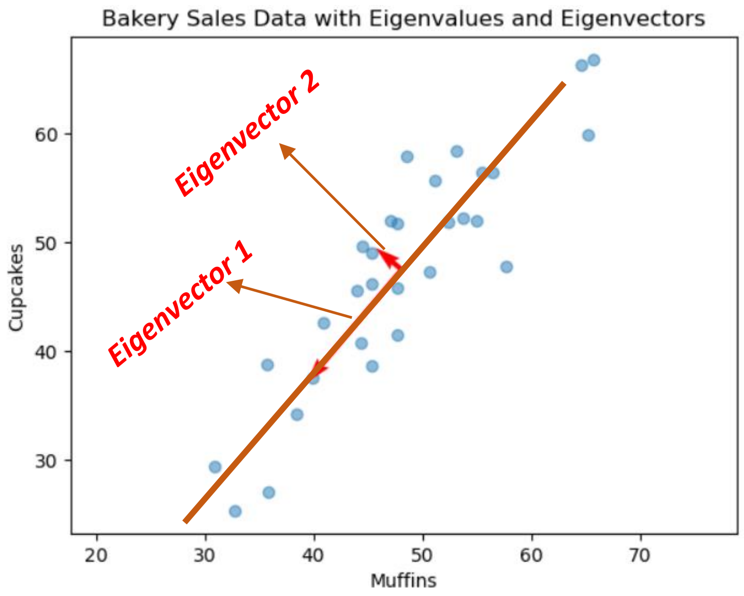



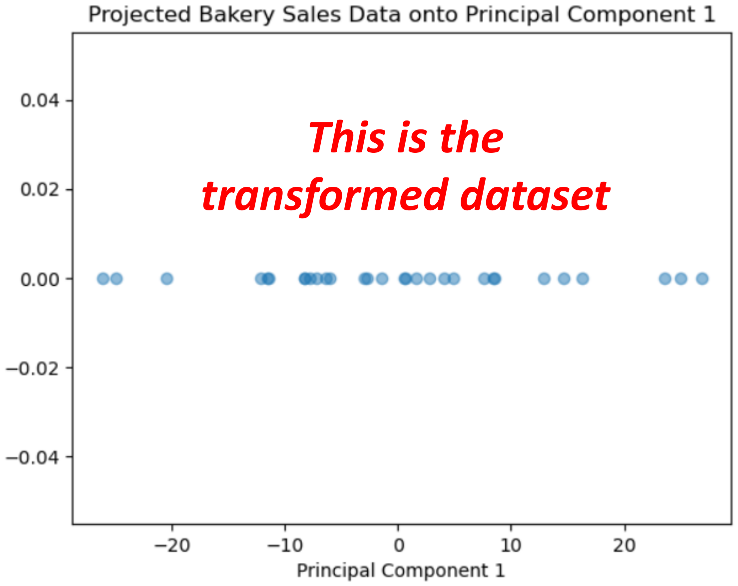
# Principle Component Analysis

- Therefore, we went from 2 dimensions to one dimension.
- This is why PCA is a dimensionality reduction algorithm

**2 Dimensions**




**1 Dimensions**



# Principle Component Analysis

- **Standardization**: Given a dataset represented by matrix X, create standardized matrix Z by subtracting the mean and dividing by the standard deviation for each variable.
- **Covariance Matrix**: Compute covariance matrix C of Z, capturing how each pair of variables in the dataset co-vary.
- **Eigenvectors and Eigenvalues**: Compute eigenvectors and eigenvalues of covariance matrix C.
- **Principal Components**: Eigenvectors define our PCs, each a direction in the original variable space where the original data vary the most.
- **Variance Explained**: Eigenvalues indicate the variance explained by each PC, larger eigenvalue means more variance explained.
- **Projection**: Project original data Z onto PCs to get scores, which are the coordinates of original data along new PC directions, yielding a lower-dimensional representation of our data.

# Conclusion

- We have learned about the basics of Linear Algebra
- Important concepts:
  - The column and row picture of linear systems
  - Invertible/non-singular matrices.
  - Matrix factorization (LU)
  - Row echelon form (REF) and reduced row echelon form (RREF)
  - Eigenvectors and eigenvalues
  - Principle component analysis

# Thank you


Q&A

# Resources


- DeepLearning.AI. (2023). Linear Algebra for Machine Learning and Data Science. Coursera. https://www.coursera.org/learn/machine-learning-linear-algebra
- Strang, G. (2016). Introduction to linear algebra (Fifth). Wellesley-Cambridge Press.
- Strang, G. (2009). Introduction to linear algebra (Fourth). Wellesley-Cambridge Press.
- Aguilar, C. O. (Year). MATH 233 - Linear Algebra I: Lecture Notes. Department of Mathematics, SUNY Geneseo
- Kolter, Z., & Do, C. (2008). Linear Algebra Review and Reference.
- Strang, G. (Year). Zoomnotes for Linear Algebra. Wellesley-Cambridge Press.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. Publisher.