University of Chicago Professional Education

MSCA 37016

Advanced Linear Algebra for Machine Learning

Session 1

Shaddy Abado Ph.D.

About Me

About Me

Shaddy Abado Ph.D.

Email

sabado@uchicago.edu

Lecture

Time:

Tuesday @ 6 PM - 9 PM CST

Thursday @ 6 PM - 9 PM CST

ABOUT YOU

What About You?

Quizzes

Pre Class
Session 1 Module

Quiz 1
Session 1 Module | **Not available until** Aug 25 at 8:30pm | **Due** Aug 28 at 11:59pm | 10 pts

Quiz 2
Session 2 Module | **Not available until** Aug 26 at 8:30pm | **Due** Aug 28 at 11:59pm | 10 pts

Quiz 3
Session 3 Module | **Not available until** Sep 1 at 11:30pm | **Due** Sep 3 at 11:59pm | 10 pts

Quiz 4
Session 4 Module | **Not available until** Sep 8 at 8:30pm | **Due** Sep 10 at 11:59pm | 10 pts

Quiz 5
Session 5 Module | **Not available until** Sep 15 at 8:30pm | **Due** Sep 17 at 11:59pm

INTRODUCTION

What is Linear Algebra?

Line-like

Linear:

- of, relating to, resembling, or having a graph that is a line and especially a straight line
- of the first degree with respect to one or more variables

Relationships

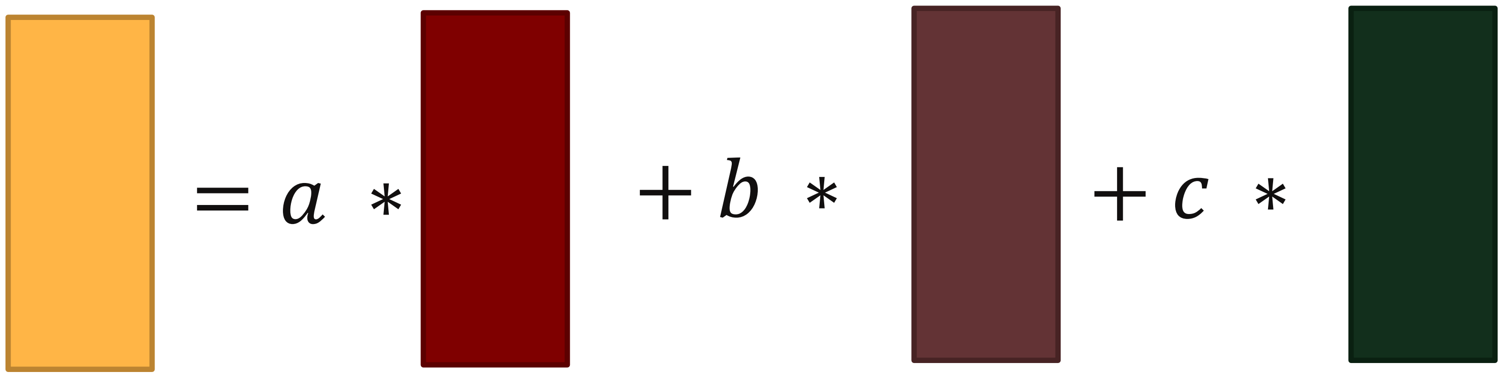
Algebra:

- From Arabic *"al-jabr"* meaning "reunion of broken parts
- A generalization of arithmetic in which letters representing numbers are combined according to the rules of arithmetic

Source: **Merriam-Webster**

Linear Algebra: “line-like relationships”

What is Linear Algebra?



Examples of linear equations

$$\left.\begin{aligned} f(x) &= a * x + b \\ f(x, y) &= a * x + y * b + c \\ f(x) &= a * x + b * x^2 + c \end{aligned}\right\} \textbf{Linear in weights}$$

What is Linear Algebra?

The fundamental problem of linear algebra is to solve n linear equations in m unknowns; for example

$$\boldsymbol{A} \cdot x = b$$

where $\boldsymbol{A}$ is a known constant rectangular matrix, b is a known column vector, and x is an unknown column vector.

What if there are more equations than unknowns?
What if there are more unknowns than equations?

WHY IS MSCA 32010 - LINEAR ALGEBRA AND MATRIX ANALYSIS LINEAR PART OF THIS PROGRAM

THE UNIVERSITY OF CHICAGO

Linear Algebra for Predictive Analytics

Data (Raw Material) → **Information** → **Knowledge** → **Action**

- Vectors and matrices allow us to represent data and understand the world.
- Linear algebra is an essential tool to transform data into knowledge and action.

From Scalar to Matrix – Car Data Example

Single Observations

Max Speed = 100 MPH

Price = $30K

Number of Doors = 4 Doors

Scalars

(Attribute of a car)

Single Car

Car ID	Speed	Price	# Doors
A	100	30	4

Vector

(A Car)

Car Fleet

Car ID	Speed	Price	# Doors
A	100	30	4
B	140	50	2
C	120	25	2
D	90	35	4

Matrix

(Car fleet)

Car Data - Price(Speed, #Doors)

Linear Equation $\boxed{\boldsymbol{Price = a * Speed + b * \#Doors + c}}$

Linear Model

$$Ax = b$$

A – Matrix (speed and # of Doors)

b - Price

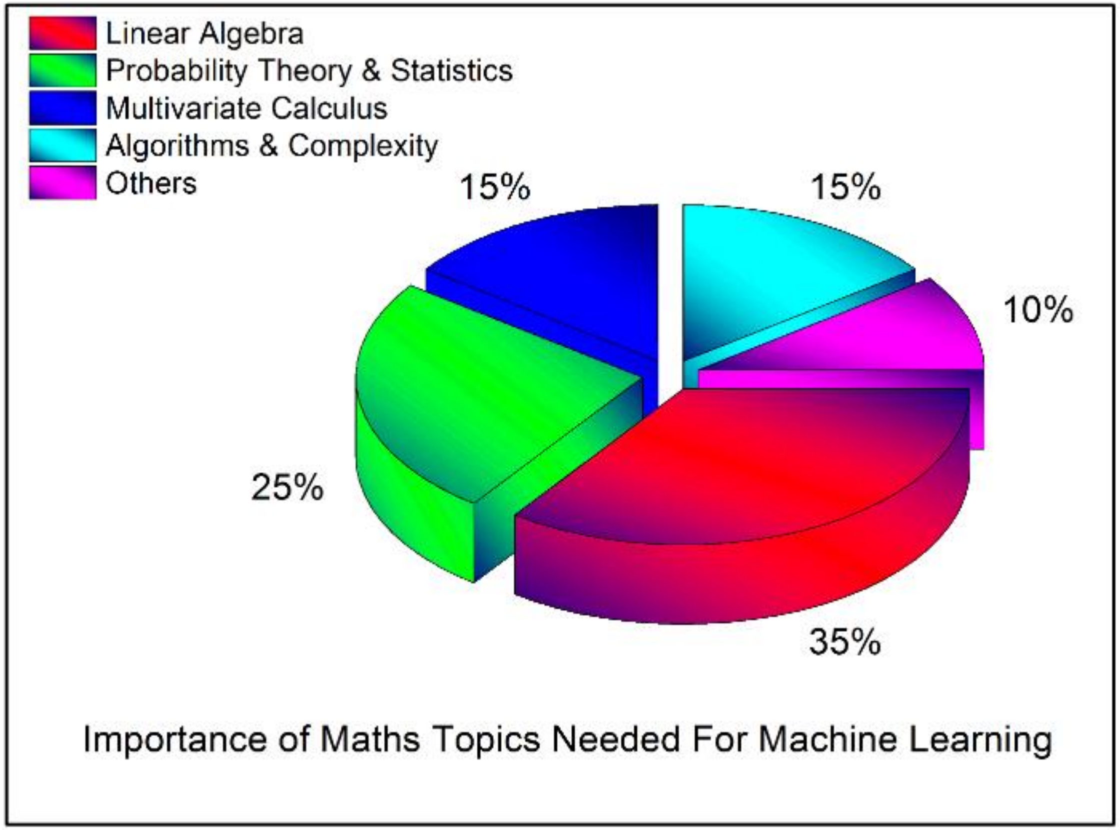
X – Coefficients (a, b and c)

Linear Algebra

Car ID	Speed	# Doors	Price ($K)
A	100	4	30
B	140	2	50
C	120	2	25
D	90	4	35

Data ➔ Information ➔ Knowledge ➔ Action

The Mathematics of Machine Learning



Principal Component Analysis (PCA), Singular Value Decomposition (SVD), Eigendecomposition of a matrix, LU Decomposition, QR Decomposition/Factorization, Symmetric Matrices, Orthogonalization & Orthonormalization, Matrix Operations, Projections, Eigenvalues & Eigenvectors, Vector Spaces and Norms

Source: http://datascience.ibm.com/blog/the-mathematics-of-machine-learning

Building Blocks – From Data To Knowledge using Linear Algebra

Data → Information → Knowledge → Action

Information and Knowledge

PCA
SVM
LSE
Logistic regression
Text mining

Neural Networks
Deep learning
Kernel Regression
etc.

Linear Algebra

Numbers (Whole, Integers, Rational, Irrational, Real, etc.)	Vectors (Norm, Space, Bases, etc.)	Matrices (Determinant, inverse, etc.)

Data

SYLLABUS

Textbook

List of recommended textbooks

- “Introduction to Linear Algebra.” Gilbert Strang; 5th Edition (2016)
- “Matrix Methods in Data Mining and Pattern Recognition.” Lars Eldén (2007)
- “Linear Algebra Tools for Data Mining.” Dan A. Simovici (2012)
- "Linear Algebra and Learning from Data" Gilbert Strang; 5th Edition (2019)
- “Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares” Stephen Boyd and Lieven Vandenberghe 1st Edition (2018)

Software

- Python 3
- **Main libraries**

  Scipy and Numpy
- **Recommended**

  Install Anaconda (https://www.anaconda.com/products/individual)

Anaconda Installers

Windows

Python 3.8

64-Bit Graphical Installer (466 MB)

32-Bit Graphical Installer (397 MB)

MacOS

Python 3.8

64-Bit Graphical Installer (462 MB)

64-Bit Command Line Installer (454 MB)

Linux

Python 3.8

64-Bit (x86) Installer (550 MB)

64-Bit (Power8 and Power9) Installer (290 MB)

TA and Grader

Section 1

Gina Champion

Email: gachampion@uchicago.edu

Section 6

Joshua Goldberg

Email: joshgoldberg@uchicago.edu

- Office hours:
  - TBD
  - office hours will be hosted online via Zoom.

Canvas

MSCA 37016 6 (Autumn 2022) Advanced Linear Algebra for Machine Learning

Edit

THE UNIVERSITY OF CHICAGO

Master of Science in Analytics

Welcome to Advanced Linear Algebra for Machine Learning. Your instructor is Shaddy Abado, whom you can reach at sabado@uchicago.edu.

Important resources are given below, including the full syllabus and links to orientations for Canvas and the course. The course schedule is below as well.

https://canvas.uchicago.edu/

Prerequisites

Undergraduate level linear algebra:

- Solving linear system of equations
- Knowledge of vector dot product and matrix multiplication
- Knowledge of key matrix operations (e.g., transpose, inverse, etc.)
- Knowledge of key types of matrices (e.g., identity, symmetric, etc.)

Evaluation

Class quizzes: 4 X 10% = 40%

Assignments: 4 X 15% = 60%

<u>**Total**: 100%</u>

Grading Scale

Pass/Fail

Pass 80-100

Fail 0-79

Assignments

- Four Assignments
- You will be asked to solve theoretical problem sets, in addition to verifying your answers using python.
- You will be asked to show and explain your work (*If you can't explain it, you don't understand it*).
- The assignments are due at the beginning of the next session and should be submitted in Canvas.

Class schedule (Section 1)

Sessions:

- Session #1 – Tuesday 8/23/2022 6pm-9pm CST
- Session #2 – Friday 8/26/2022 6pm-9pm CST
- Session #3 – Tuesday 8/30/2022 6pm-9pm CST
- Session #4 – Tuesday 9/6/2022 6pm-9pm CST
- Session #5 – Tuesday 9/13/2022 6pm-9pm CST

Quizzes:

- Quiz #1 - Due Thursday 8/25/2022 Midnight CST
- Quiz #2 - Due Sunday 8/28/2022 Midnight CST
- Quiz #3 - Due Thursday 9/1/2022 Midnight CST
- Quiz #4 - Due Thursday 9/8/2022 Midnight CST

Assignments:

- Assignment #1 - Due Tuesday 8/30/2022 5:59pm CST
- Assignment #2 - Due Sunday 9/4/2022 5:59pm CST
- Assignment #3 - Due Tuesday 9/6/2022 5:59pm CST
- Assignment #4 - Due Saturday 9/10/2022 Midnight CST

Class schedule (Section 6)

Sessions:

- Session #1 – Thursday 8/25/2022 9 6pm-9pm CST
- Session #2 – Friday 8/26/2022 6pm-9pm CST
- Session #3 – Thursday 9/1/2022 6pm-9pm CST
- Session #4 – Thursday 9/8/2022 6pm-9pm CST
- Session #5 – Thursday 9/15/2022 6pm-9pm CST

Quizzes:

- Quiz #1 - Due Saturday 8/28/2022 Midnight CST
- Quiz #2 - Due Sunday 8/28/2022 Midnight CST
- Quiz #3 - Due Saturday 9/3/2022 Midnight CST
- Quiz #4 - Due Saturday 9/10/2022 Midnight CST

Assignments:

- Assignment #1 - Due Thursday 9/1/2022 5:59pm CST
- Assignment #2 - Due Sunday 9/4/2022 5:59pm CST
- Assignment #3 - Due Thursday 9/8/2022 5:59pm CST
- Assignment #4 - Due Saturday 9/10/2022 Midnight CST

Course Schedule

1. Introduction
2. Solving Linear Equation and Vector Spaces
3. Least Squares Approximation and Linear Transformation
4. Eigenvalues, Eigenvectors and Singular Value Decomposition
5. Dynamic matrices and Tensor Math

Important Note: Changes may occur to the syllabus at the instructor's discretion. When changes are made, students will be notified

Late work and Attendance Policies

Late work Policy

All assignments must be submitted to this course's Canvas site. Late assignments are not accepted without explicit permission from the instructor, and permission can be granted only in the case of an emergency and prior to the assignment due date. These extensions should be approved at least 24 hours prior to any deadline. Late work will be subject to a loss of 50% of the assigned grade. You are asked to plan your time accordingly to avoid any late submissions.

Attendance Policy

This course will meet weekly. Your attendance at all 5 sessions is required and paramount to your success in this class.

NOTATIONS AND CONVENTIONS

THE UNIVERSITY OF CHICAGO

Notations and Conventions

- Scalar
  - a, b, c, etc. or λ, ρ, etc.
- Vector
  - $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$, etc. or $\vec{u}$, $\vec{v}$, $\vec{w}$, etc.
- Matrix
  - $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{Q}$, etc.
- Scalar Product (dot product, inner product)

$$\bullet$$

- Scalar Multiplication

$$*$$

Scalars

Types of Numbers

Real numbers often result from making measurements and measurements are always approximate

Real Numbers $\mathbb{R}$: Rational + Irrational

Irrational $\mathbb{I}$: can't be represented as a/b where a and b are integers. For example, π = 3.14159..., and $\sqrt{2}$ = 1.4142....

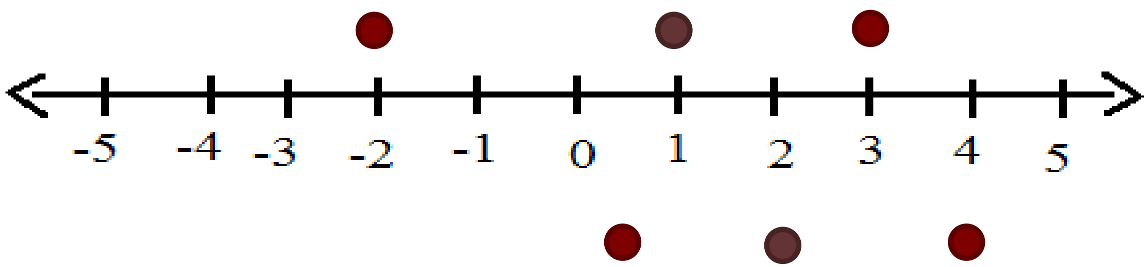
Rational $\mathbb{Q}$: a/b where a and b are integer 2/1, -1/2, 2/5 etc.

Integers $\mathbb{Z}$: .... -3, -2, -1, 0, 1, 2, 3, ...

Natural $\mathbb{N}$: 1, 2, 3, ...
And
Whole: 0,1,2,3, ...

Scalars

Definition:
Quantity having only magnitude, not direction.



$$-2 + 3 = 1$$

$$1/2 * 4 = 2$$

INTRODUCTION TO VECTORS

THE UNIVERSITY OF CHICAGO

So again, what is Linear Algebra?

Linear algebra is the study of vectors and linear functions

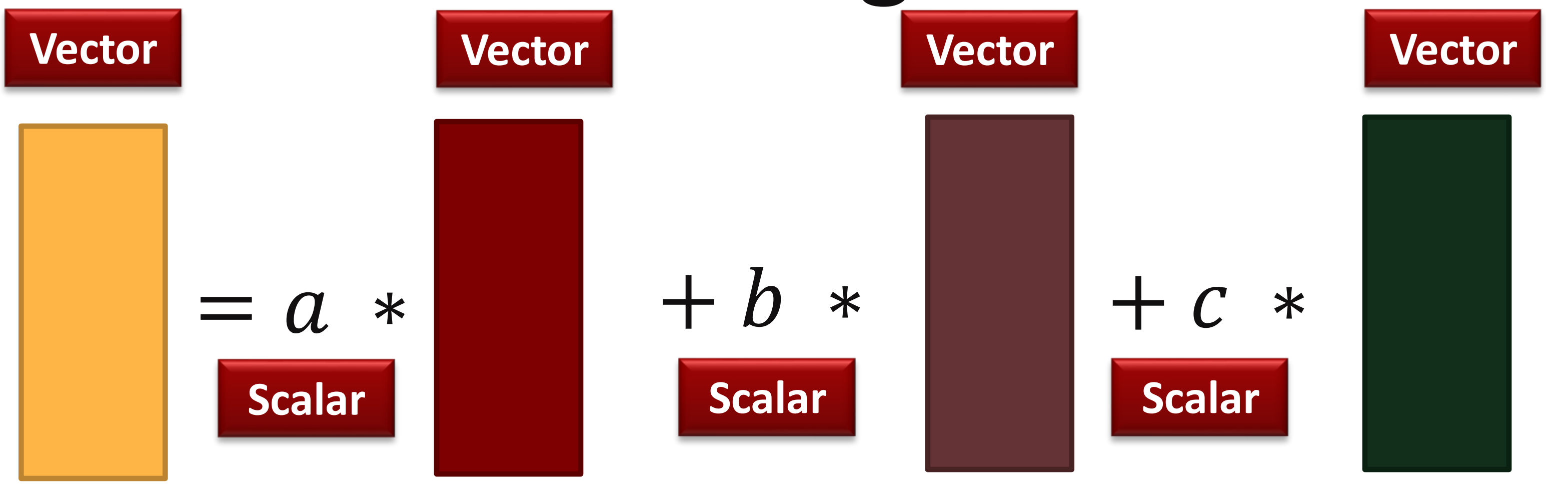
- What are vectors?
- What are linear functions?

In broad terms, vectors are things you can

(1) add and (2) scalar multiply

Linear functions are functions of vectors that respect these properties $T(c\boldsymbol{v} + d\boldsymbol{w}) = cT(\boldsymbol{v}) + dT(\boldsymbol{w})$

What is Linear Algebra?



Linear Functions

Examples of linear functions

$$f(x) = a * x + b$$

$$f(x, y) = a * x + y * b + c$$

$$f(x) = a * x + b * x^2 + c$$

Linear in weights

What are vectors?

Vectors are things you can add and scalar multiply.

For example, vectors can be:

Numbers: $2 * 1 + 1$

Polynomials: $3 * x + 4 * x^2$

Functions: $1.5 * \sin x + 2 * \cos x$

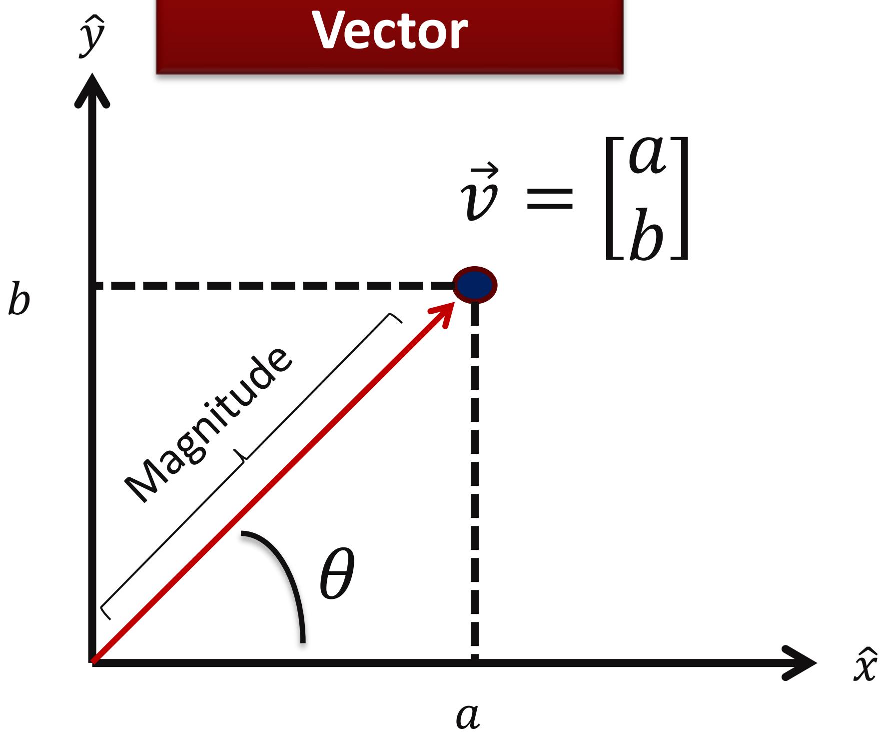
Two column vectors: $\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 6 \\ 0.4 \end{bmatrix}$

Vectors – Two and Three Dimensional

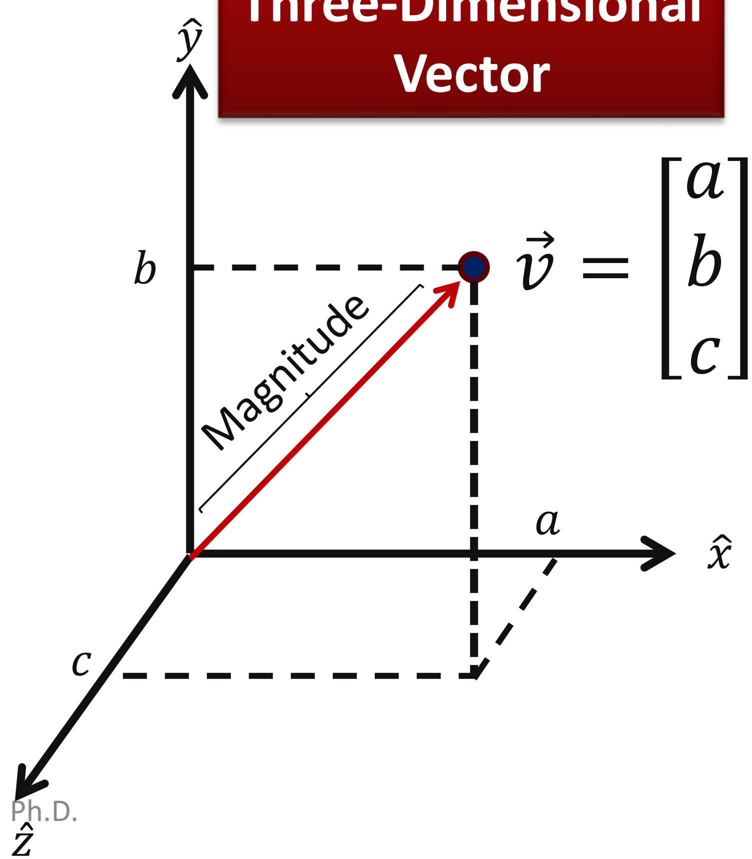
Vector is a quantity that has magnitude and direction

For simplicity, we will restrict our discussion to Euclidean space

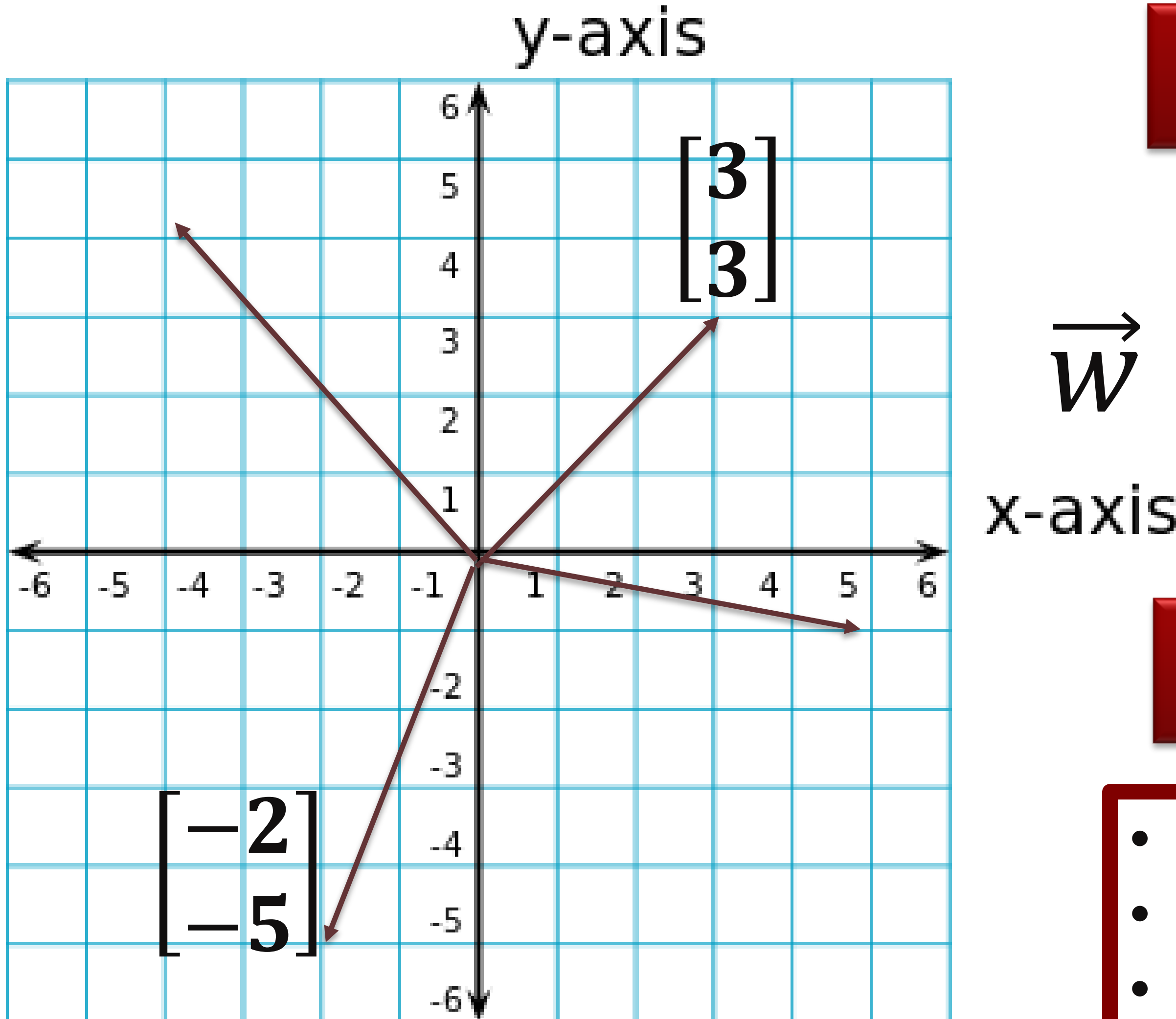
Two-Dimensional Vector



Three-Dimensional Vector



Two Dimensional – Vector Representation and Visualization



X-axis coordinate (scalar)

$$\vec{w} = \begin{bmatrix} x \\ y \end{bmatrix}$$

Y-axis coordinate (scalar)

- **Two Numbers**
- **Arrow from [0, 0]**
- **Point in space**

Notations and Conventions

In this course, we will be defining vectors in column notations. However, row notation may also be used for convenience.

$$\vec{w} = \begin{bmatrix} x \\ y \end{bmatrix} = (x, y)$$

We will be balancing between algebraic and geometrical notation.

Definition:

Vector Addition/Subtraction

$$\vec{v} + \vec{w} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

1st and 2nd components of w

$$= \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix}$$

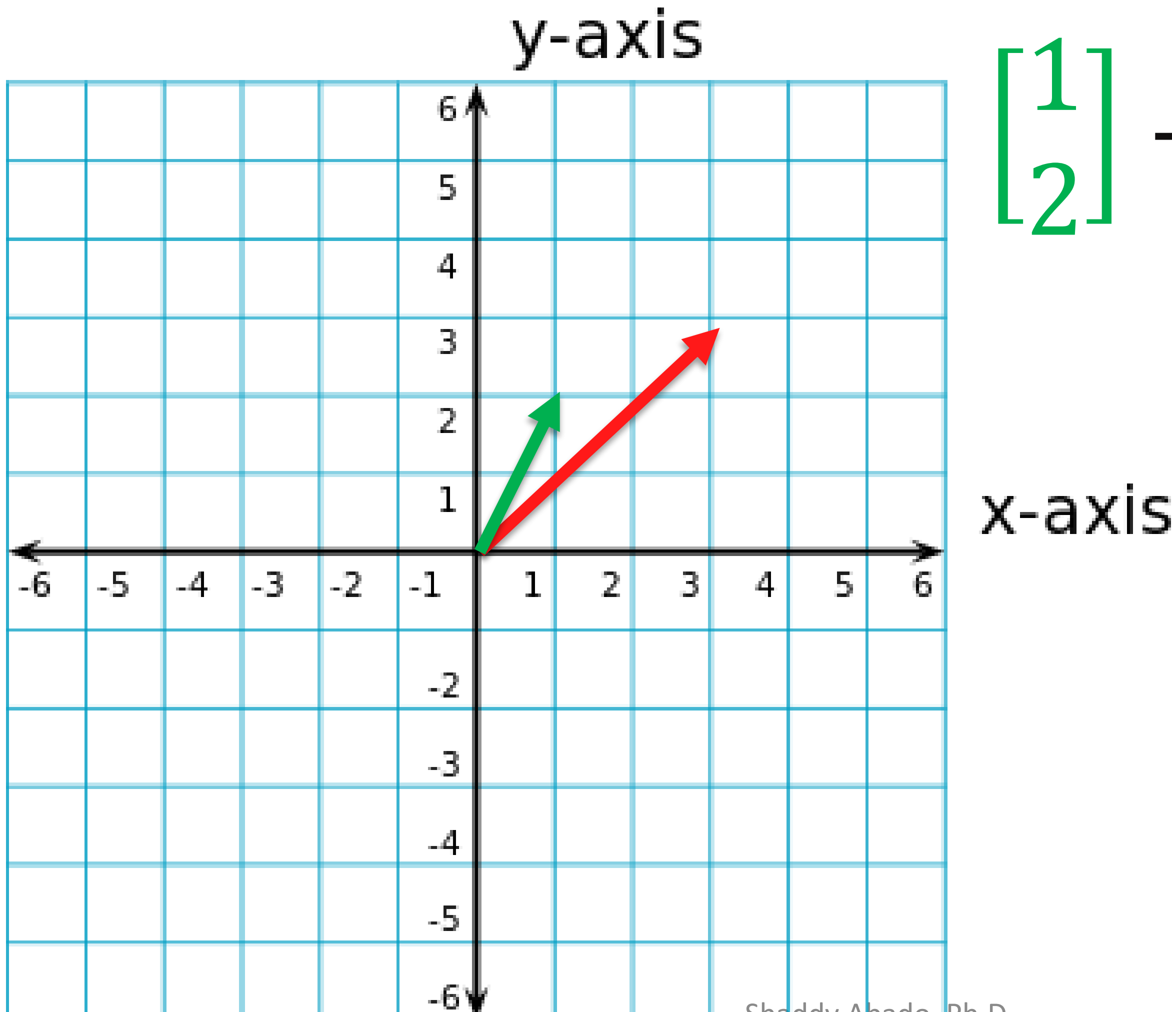
Addition

$$\vec{v} - \vec{w} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} - \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$= \begin{bmatrix} v_1 - w_1 \\ v_2 - w_2 \end{bmatrix}$$

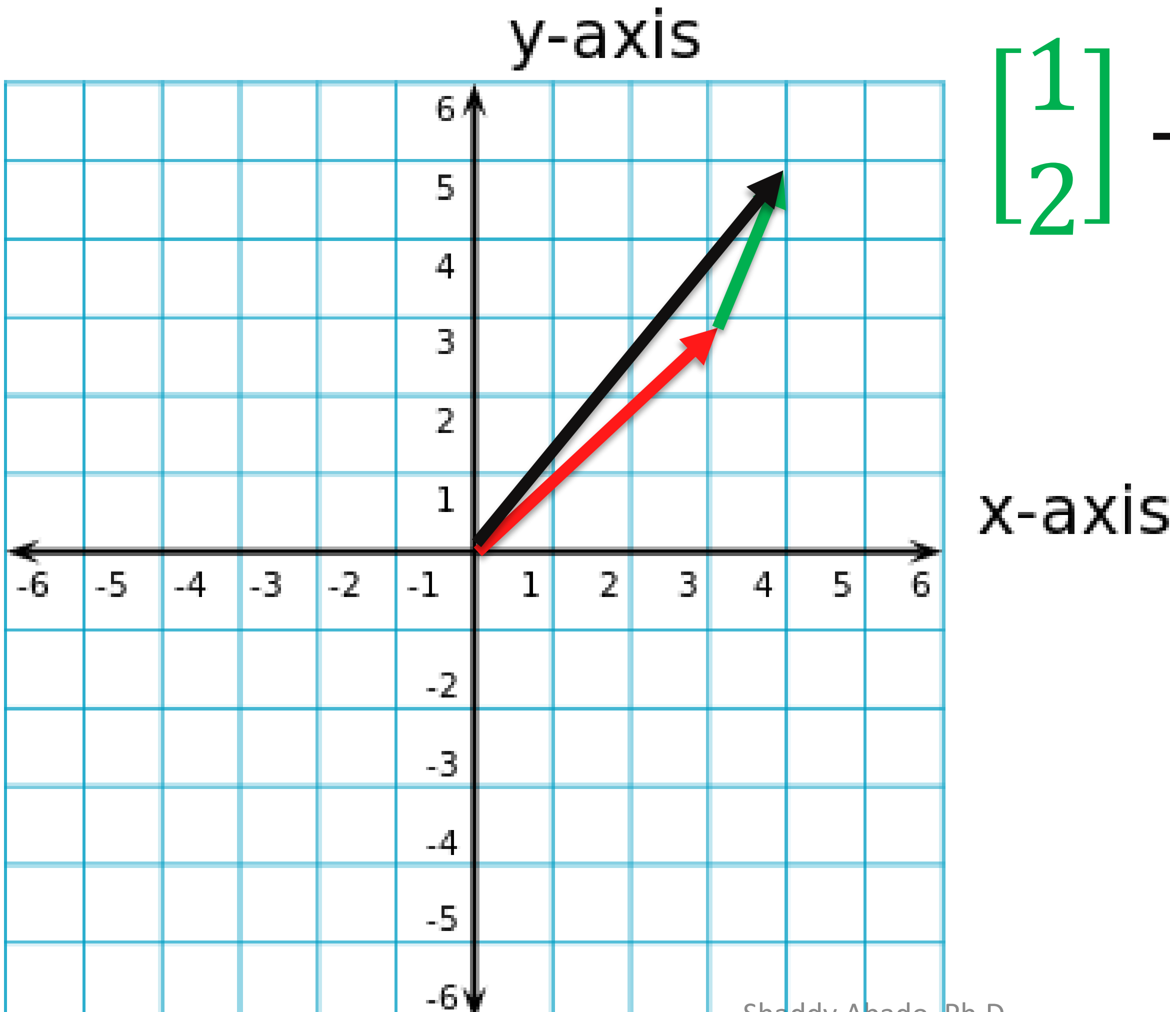
Subtraction

Vector Addition - Example



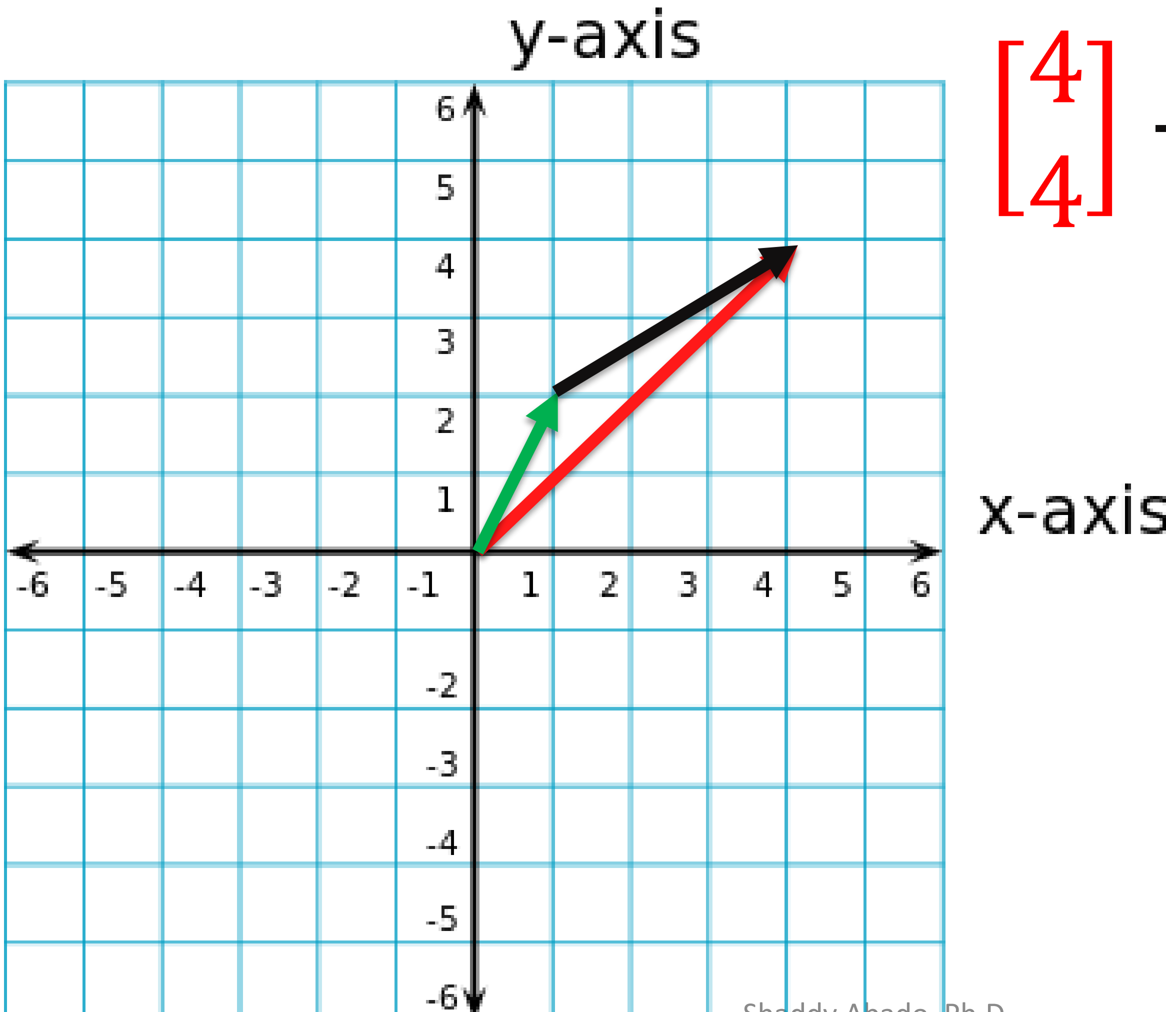
$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

Vector Addition - Example



$$\begin{bmatrix}1\\2\end{bmatrix} + \begin{bmatrix}3\\3\end{bmatrix} = \begin{bmatrix}4\\5\end{bmatrix}$$

Vector Subtraction - Example



$$\begin{bmatrix}4\\4\end{bmatrix} - \begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}3\\2\end{bmatrix}$$

Definition:

Zero vector

$$\begin{bmatrix}1\\2\end{bmatrix} - \begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} = \vec{0}$$

A vector with zero components

Null Vector
Or
Zero vector

- Is the zero vector the same as 0? (zero scalar)
- What about $0 * x$? (zero function)

Vector Addition/Subtraction

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 0 \\ -10 \end{bmatrix} = ?$$

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} - 3 = ?$$

Vectors of different 'shapes' can't be added

Definition:

Scalar Multiplication

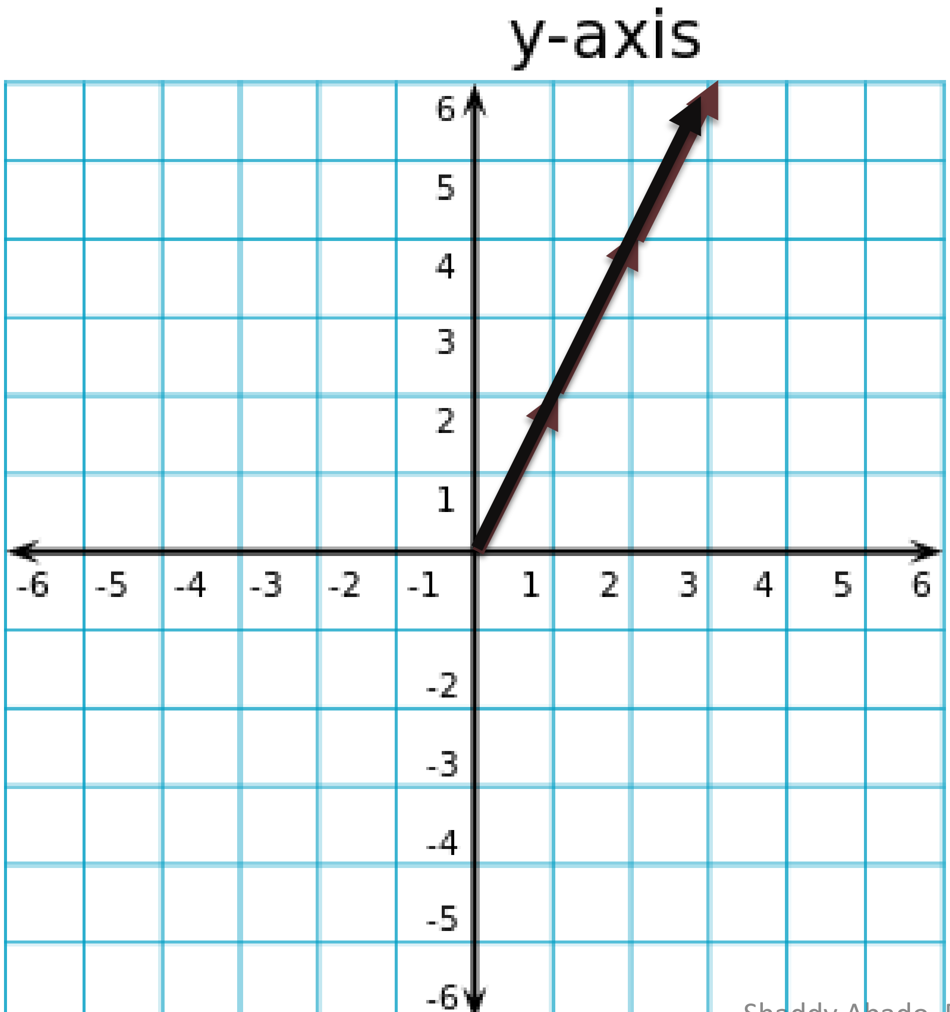
Scalar multiplication changes the **scale** of the arrow from the origin to the point

$$b * \vec{v} = b\vec{v} = \overbrace{\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \cdots + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}}^{\textbf{Add vector } v \text{ } b\textbf{-times}}$$

$$= \begin{bmatrix} v_1 + \cdots + v_1 \\ v_2 + \cdots + v_2 \end{bmatrix} = \begin{bmatrix} b * v_1 \\ b * v_2 \end{bmatrix}$$

Example:

Scalar Multiplication

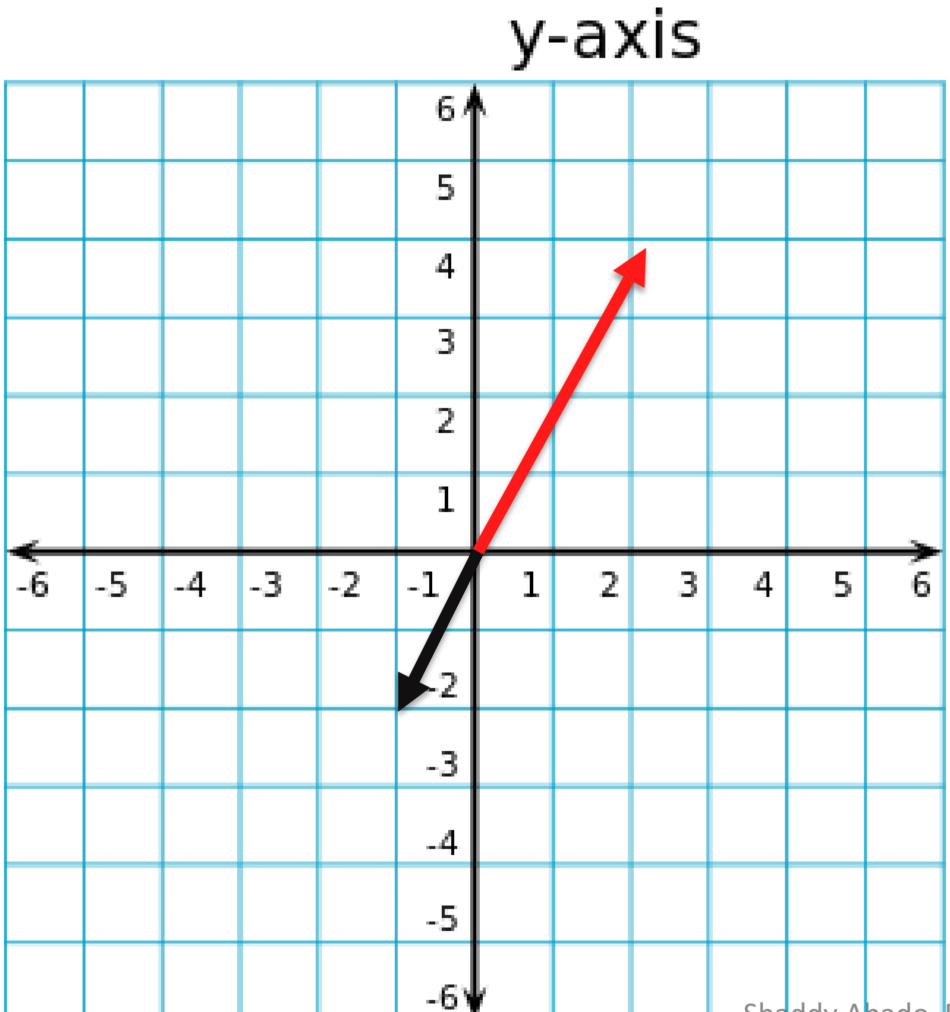


$$3 * \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 * 1 \\ 3 * 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Scalar Multiplication Changes the **scale** of the arrow from the origin to the point (The angle w.r.t to x and y-axes is unchanged)

Example:

Scalar Multiplication



Reflection

$$-0.5 * \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

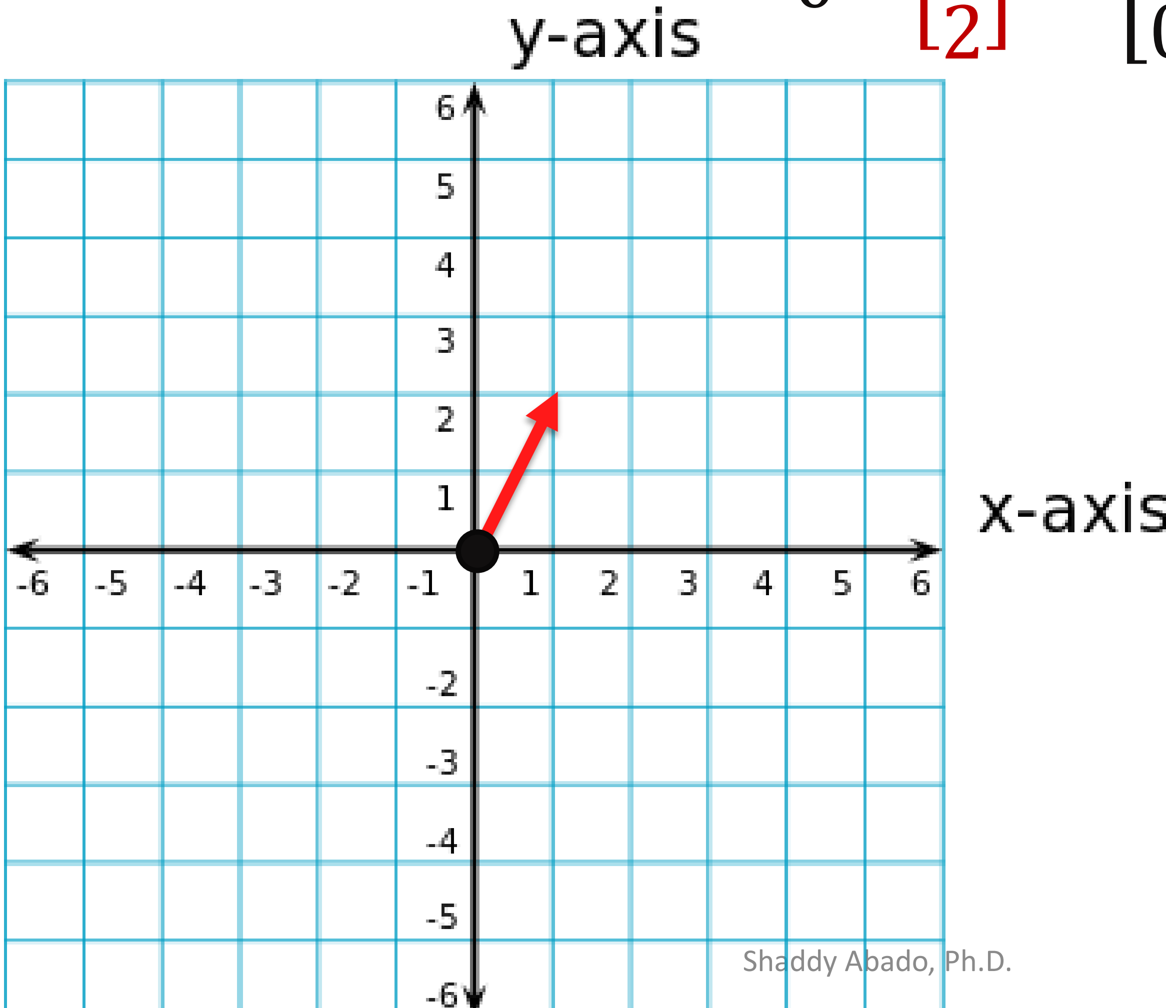
$$= \begin{bmatrix} -0.5 * 2 \\ -0.5 * 4 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

Scalar Multiplication Changes the **scale** of the arrow from the origin to the point (The angle w.r.t to x and y-axes is unchanged)

Example:

Scalar Multiplication

$$0 * \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 * 1 \\ 0 * 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}$$



Null Vector
Or
Zero vector

Break

THE UNIVERSITY OF CHICAGO

Fundamental Concept I:
Linear Combination

The sum of $a\vec{v} + \; b\overrightarrow{w}$ *is a* ***linear combination***

Where

$\vec{v}$ and $\overrightarrow{w}$ are vectors

a and b are scalars

Four unique cases:

$$\boldsymbol{Sum}\text{: } 1\vec{v} + 1\overrightarrow{w} = \vec{v} + \overrightarrow{w}$$

$$\boldsymbol{Difference}\text{: } 1\vec{v} - 1\overrightarrow{w} = \vec{v} - \overrightarrow{w}$$

$$\boldsymbol{Zero}\text{: } 0\vec{v} + 0\overrightarrow{w} = \overrightarrow{\mathbf{0}}$$

$$\boldsymbol{Scalar\ Multiplication}\text{: } c\vec{v} + 0\overrightarrow{w} = c\vec{v}$$

Example: Linear combination

$$-0.5 * \begin{bmatrix} 2 \\ 4 \end{bmatrix} + 2 * \begin{bmatrix} 1 \\ 5 \end{bmatrix} =$$

$$= \begin{bmatrix} -0.5 * 2 \\ -0.5 * 4 \end{bmatrix} + \begin{bmatrix} 2 * 1 \\ 2 * 5 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 10 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

Fundamental Concept II:

Independency and Dependency

$$a\vec{v} + b\vec{u} = \vec{0}$$

- Vectors v and u are **Independent** if no combination except $0\vec{v} + 0\vec{u}$ gives $\vec{0}$
- Vectors v and u are **Dependent** if there is a combination $a\vec{v} + b\vec{u}$ that gives $\vec{0}$

$$\begin{bmatrix}1\\0\end{bmatrix} \ and \begin{bmatrix}0\\1\end{bmatrix}$$

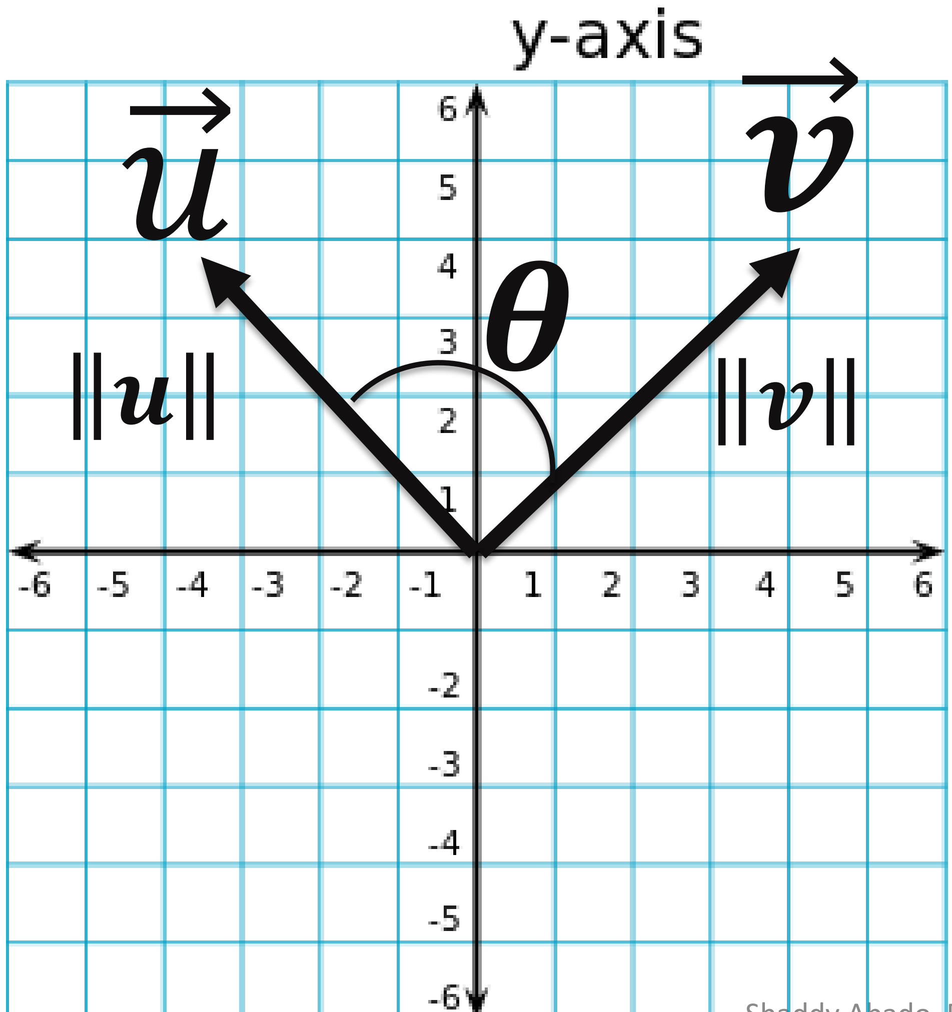
Independent

$$\begin{bmatrix}1\\1\end{bmatrix} \ and \begin{bmatrix}2\\2\end{bmatrix}$$

Dependent

Next:

Vector Norm and Angle between two vectors



✓ Addition/Subtraction

✓ Scalar Multiplication

Next …

- Length/Magnitude of vectors?
- Angle between vectors?

θ **Direction**

$||v||$ **Magnitude**

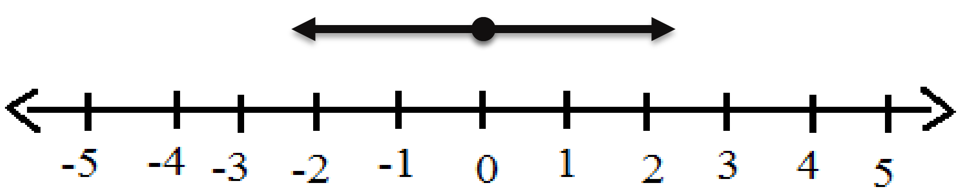
VECTOR NORM

Vector Norm: Absolute Value and Vector Length

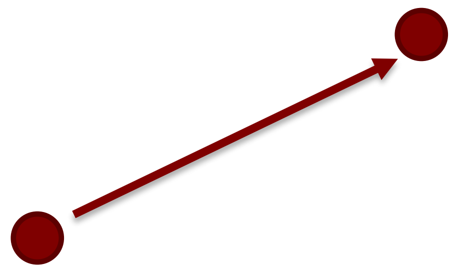
Definition

Absolute value is the size of scalar - $|a|$

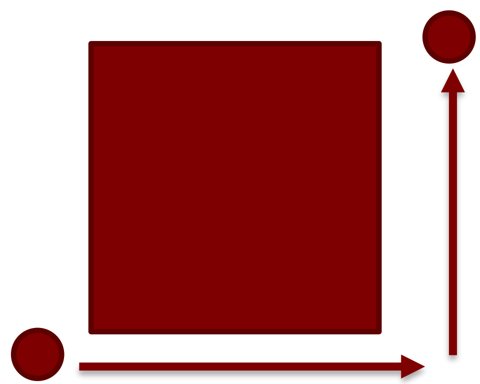
$$|-2| = 2$$
$$|2| = 2$$



Vector Norm: Absolute Value and Vector Length



$L_2\ Norm$ –
Euclidian Distance



$L_1\ Norm$ –
Manhattan Distance

$L_\infty\ Norm$

Definitions:

Vector Norm and $L_p\ Norm\ (p - norm)$

Definition

Vector Norm measures the magnitude of a vector

Definition

Let $p \geq 1$ be a real number. The p-norm of vector $\vec{v}$ is:

$$\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix}$$

$$\|\vec{\boldsymbol{v}}\|_p = \sqrt[p]{|v_1|^p + |v_2|^p + \cdots + |v_N|^p}$$

Definition:

$$L_1, L_2, and\ L_\infty, Norms$$

- $L_1\ Norm$ - Manhattan Distance
- $L_2\ Norm$ – **Euclidian Distance**
- $L_\infty\ Norm$

Note:

$L_2\ Norm$ is a generalization of the standard Euclidian distance in two dimensional to N-dimensional

L_2 $Norm$ (Euclidian Distance)

$$\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix}$$

Definition

$$\|\vec{v}\|_2 = \sqrt{v_1^2 + v_2^2 + \cdots v_N^2}$$

Examples

$$\left\|\begin{matrix} 2 \\ 3 \end{matrix}\right\|_2 = \sqrt{2^2 + 3^2} = \sqrt{4 + 9} = \sqrt{13}$$

$$\left\|\begin{matrix} 1 \\ 2 \\ 3 \end{matrix}\right\|_2 = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{1 + 4 + 9} = \sqrt{14}$$

Vector Norm Properties

The *norm* $||v||$ of a vector $v \in S$ is a real number that satisfies the following properties:

- $||v|| \geq 0$
- $||v|| = 0$ if and only if $v = 0$,
- $||av|| = |a|\ ||v||$, a $\in R_1$, and
- $||v + w|| \leq ||v|| + ||w||$, (triangle or Minkowski inequality).

Unit Vector

A unit vector u is a vector whose length equals one → $\|u\| = 1$

$$u = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\|u\| = \sqrt{0^2 + 1^2} = \sqrt{0+1} = \sqrt{1} = 1$$

What about the following vectors?

$$\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{bmatrix}$$

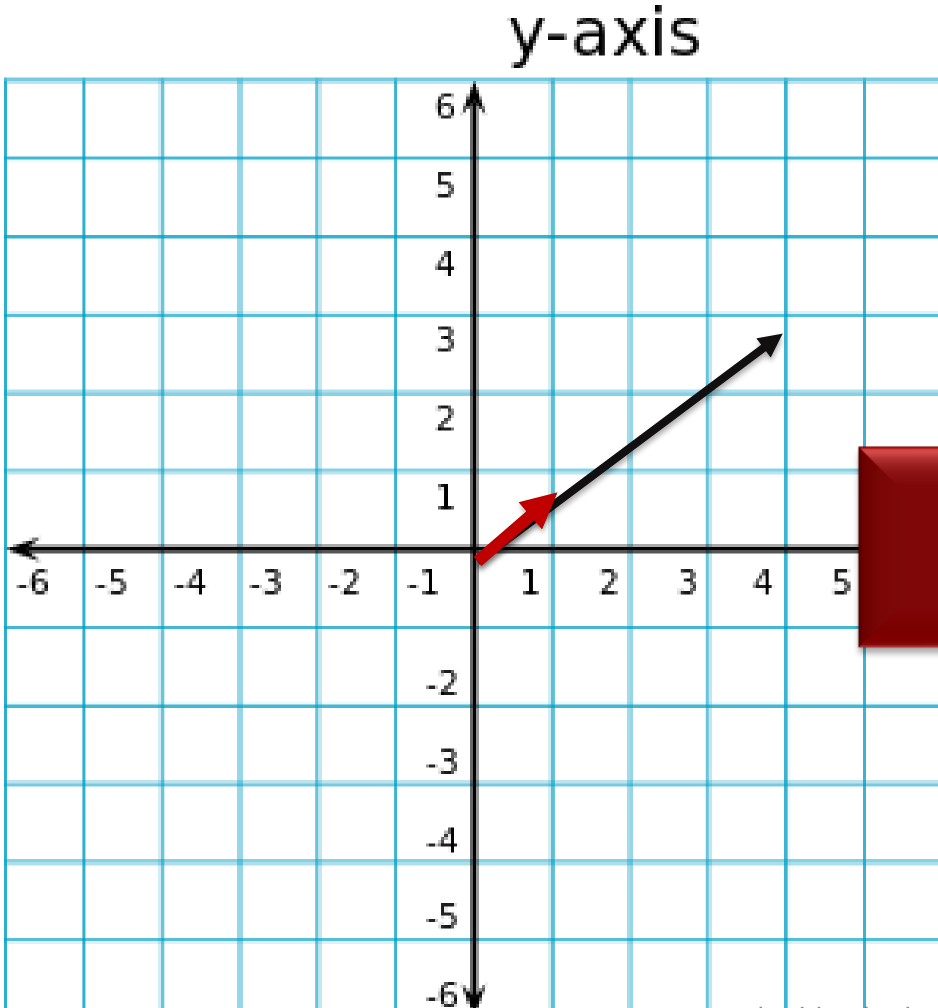
Check

$$\|u\| = \sqrt{(1/\sqrt{3})^2 + (\sqrt{2}/\sqrt{3})^2} = \sqrt{1/3 + 2/3} = \sqrt{1} = 1$$

Vector Normalization

$$\boldsymbol{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

$$\|\boldsymbol{v}\| = 5$$



$$\boldsymbol{u} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}$$

This is a unit vector in the same direction as v

$$\boldsymbol{u} = \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}; \quad \|\boldsymbol{u}\| = 1$$

Check

SCALAR PRODUCT (DOT PRODUCT, INNER PRODUCT

Definition:

Scalar product (Dot product, Inner product)

It is called scalar product because the output is a scalar

$$\left\{\begin{aligned} y &= \vec{v} \cdot \vec{w} = \sum_{i=1}^{N} v_i w_i \\ &= v_1 w_1 + v_2 w_2 + \cdots v_N w_N \end{aligned}\right.$$

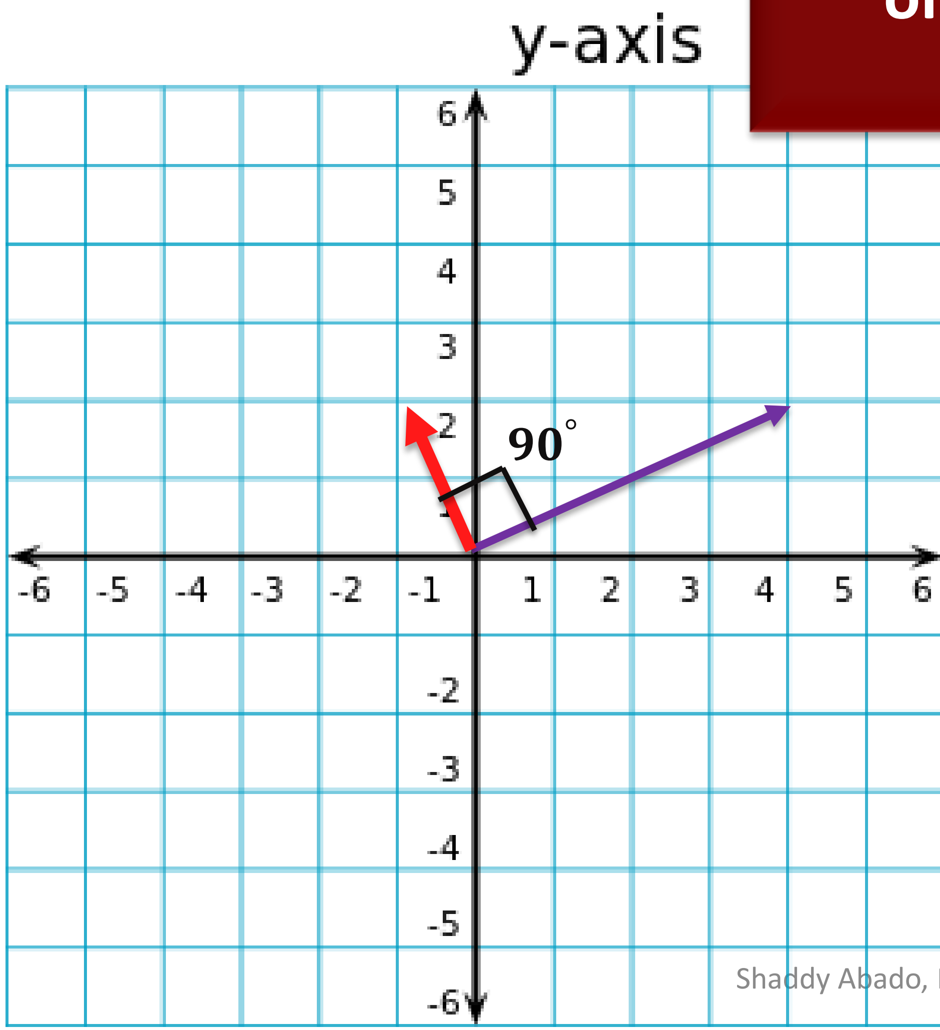
$$v = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad u = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$$

$$\left\{\begin{aligned} &\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 3 \end{bmatrix} \\ &= 1 * 7 + 2 * 3 = 13 \end{aligned}\right.$$

Definition: Orthogonal Vectors

Vectors v and u are said to be orthogonal to each other if

$$v \cdot u = 0$$



$$\begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$= 4 * -1 + 2 * 2$$

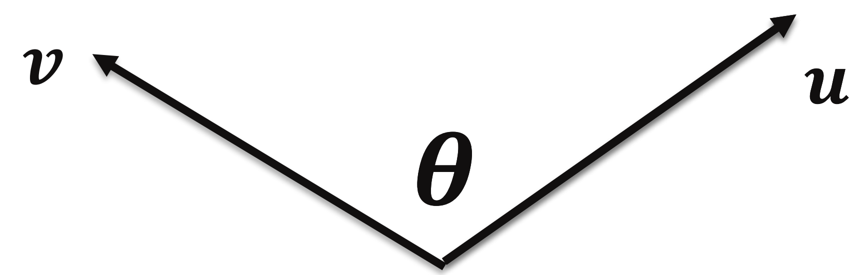
$$= 0$$

Right angle (90°) between vectors

Next: Angle Between Two Vectors

- ✓ Length
- ✓ Inner product

What about the angle $\boldsymbol{\theta}$ between vectors?



Cosine Formula:

$$\frac{\boldsymbol{v} \cdot \boldsymbol{u}}{\|\boldsymbol{v}\|\|\boldsymbol{u}\|} = \cos\boldsymbol{\theta}$$

Definition:

Cosine Formula

Schwarz's Inequality
$$|v \cdot u| \leq \|v\|\|u\|$$

$$\frac{\boldsymbol{v} \cdot \boldsymbol{u}}{\|\boldsymbol{v}\|\|\boldsymbol{u}\|} = \cos \boldsymbol{\theta}$$

$$\underbrace{\boldsymbol{v} \cdot \boldsymbol{u}}_{\textbf{Algebraic}} = \underbrace{\|\boldsymbol{v}\|\|\boldsymbol{u}\| \cos \boldsymbol{\theta}}_{\textbf{Geometric}}$$

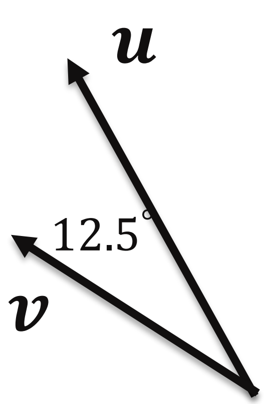
The significance of this property is that the left-hand side is purely <u>algebraic,</u> and the right-hand side is purely <u>geometric</u>.

Example: Cosine Formula

$$\frac{\boldsymbol{v} \cdot \boldsymbol{u}}{\|\boldsymbol{v}\|\|\boldsymbol{u}\|} = \mathrm{Cos}\,\boldsymbol{\theta}$$

$$\boldsymbol{v} \cdot \boldsymbol{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 4 \end{bmatrix} = -1 * -1 + 2 * 4 = 9$$

$$\|\boldsymbol{v}\| = \sqrt{1+4} = \sqrt{5} \qquad \|\boldsymbol{u}\| = \sqrt{1+16} = \sqrt{17}$$



$$\frac{9}{\sqrt{17}\sqrt{5}} = \mathrm{Cos}\,\boldsymbol{\theta}$$

$$\mathrm{Cos}\,\boldsymbol{\theta} = 0.97$$

$$\boldsymbol{\theta} \sim 0.2\ Rad\ ;\ 12.5^{\circ}$$

Check

MATRIX ALGEBRA

Matrix Algebra - Motivation

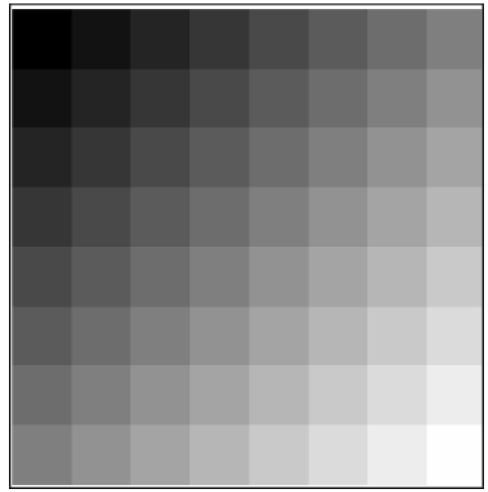
- Our ability to analyze dataset and solve equations depends on performing algebraic operations with matrices.
- Matrices are an efficient way to store information and a powerful tool for calculations involving linear transformations.
- Basic understanding of how to manipulate matrices is needed.

Keep in mind …

Matrices are the result of organizing information related to linear functions.

We are not studying matrices but rather linear functions; those linear functions can be represented as matrices under certain notational conventions.

Example I: Matrices of Images

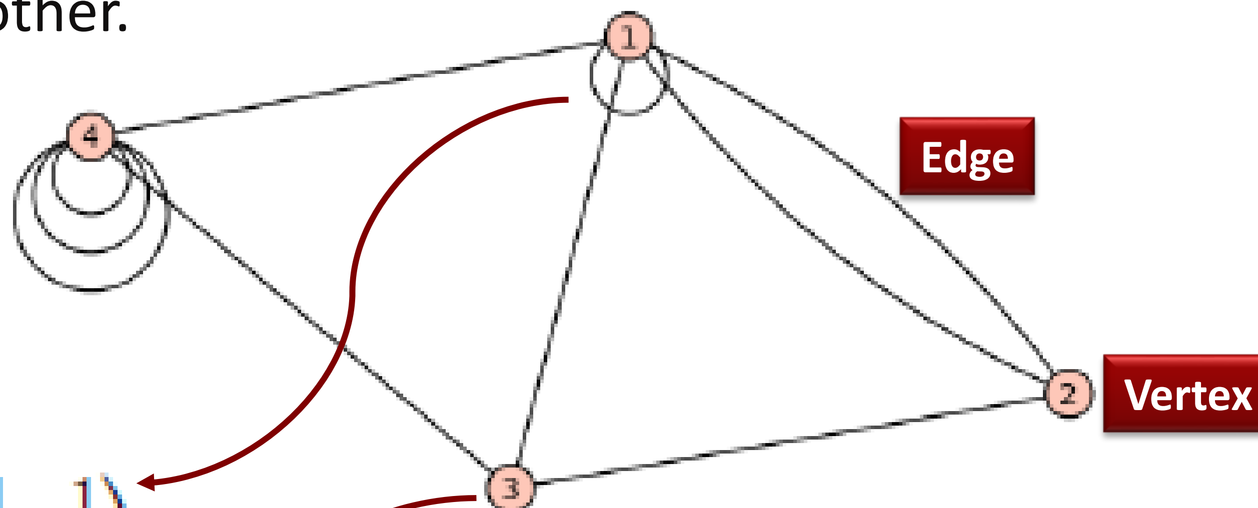


0 50 100 150 200 250 300 350 50 100
150 200 250 300 350 400 100 150 200
250 300 350 400 450 150 200 250 300
350 400 450 500 200 250 300 350 400
450 500 550 250 300 350 400 450 500
550 600 300 350 400 450 500 550 600
650 350 400 450 500 550 600 650 700

Source: "Image and its matrix, matrix and its image" by Vesna Vučković

Example II: Graph Theory

- In graph theory, a graph is a collection of vertices and some edges connecting vertices. Graphs occur in many applications, ranging from telephone networks to airline routes.
- A matrix can be used to indicate how many edges attach one vertex to another.



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

Definition:

Matrix

Scalars → Vectors → Matrices

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \qquad w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \qquad u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$$

$$\boldsymbol{A} = [v \quad w \quad u] = \begin{bmatrix} v_1 & w_1 & u_1 \\ \vdots & \vdots & \vdots \\ v_n & w_n & u_n \end{bmatrix}$$

The columns of $\boldsymbol{A}$ are vectors in $\mathbb{R}^n$

Definition:

Matrix

We will denote a matrix of size $N \times M$ as

$$\left.\overbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1M-1} & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M-1} & a_{2M} \\ & & \vdots & & \\ a_{N-1\,1} & a_{N-1\,2} & \cdots & a_{N-1\,M-1} & a_{N-1\,M} \\ a_{N1} & a_{N2} & \cdots & a_{NM-1} & a_{NM} \end{bmatrix}}^{M\ \text{Columns}}\right\} \begin{array}{c} N \\ \text{Rows} \end{array}$$

a_{ij}

i – row

j - column

Matrix Element (Scalars)

Definition:

Matrix

A vector is a special type of matrix.

- Nx1 matrix ➔ n-dimensional column vector
- 1xM matrix ➔ m-dimensional row vector

$$v_{Nx1} = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix} \qquad v_{1xM} = \begin{bmatrix} v_1 & \cdots & v_M \end{bmatrix}$$

Unless otherwise stated vectors are assumed to be column vectors.

Definition:

Matrix

Main diagonal

$$A_{3X3} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Square Matrix

$N = M$

$$A_{2X3} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Rectangular Matrix

$N \neq M$
$N > M$
Or
$N < M$

Definition:

Matrix Addition / Subtraction

Addition of matrices can be defined as

$$C_{nxm} = A_{nxm} + B_{nxm}$$

where the elements of C are obtained by adding the corresponding elements of $\boldsymbol{A}$ and $\boldsymbol{B}$.

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} b_{11} + a_{11} & b_{12} + a_{12} & b_{13} + a_{13} \\ b_{21} + a_{21} & b_{22} + a_{22} & b_{23} + a_{23} \\ b_{31} + a_{31} & b_{32} + a_{32} & b_{33} + a_{33} \end{bmatrix}$$

What about Subtraction?

Definition:

Null Matrix

$$\begin{bmatrix} 1 & 3 & -2 \\ -\frac{1}{3} & 6 & -1 \\ 2 & 2 & 5 \end{bmatrix} - \begin{bmatrix} 1 & 3 & -2 \\ -\frac{1}{3} & 6 & -1 \\ 2 & 2 & 5 \end{bmatrix} = ?$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{0}_{3X3}$$

Null Matrix
Or
Zero Matrix

Matrix Multiplication

Given matrix $\boldsymbol{A}$ we can multiply it by a:

- **Scalar**
- **Vector**
- **Matrix**

Definition:

Matrix – Scalar Multiplication

Multiplication of a matrix A by a scalar b can be defined as

$$b\boldsymbol{A}_{nXm} = \boldsymbol{C}_{nXm}$$

where the elements of $\boldsymbol{C}$ are the corresponding elements of $\boldsymbol{A}$ multiplied by b.

Matrix

Scalar

$$b * \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} =$$

$$\begin{bmatrix} b * a_{11} & b * a_{12} & b * a_{13} \\ b * a_{21} & b * a_{22} & b * a_{23} \\ b * a_{31} & b * a_{32} & b * a_{33} \end{bmatrix}$$

MATRIX – VECTOR MULTIPLICATION

Definition:

Matrix – Vector Multiplication

Multiplication of matrix $\boldsymbol{A}$ and vector $\vec{\boldsymbol{v}}$ can be defined only if they are of the proper sizes.

$$y_{mx1} = \boldsymbol{A}_{mxn} v_{nx1}$$

Note:

The number of columns of A = Number of rows of v

Definition:

Matrix – Vector Multiplication

$$y_{mx1} = A_{mxn} v_{nx1}$$

Note that # of columns of A = # of rows of v

The general formula for a matrix-vector product is

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

The vector y is a linear combination of the columns of A

Example I:

Matrix – Vector Multiplication

$$A_{2X2} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad v_{2X1} = \begin{bmatrix} 5 \\ 6 \end{bmatrix} \longrightarrow \text{"Row Weights"}$$

First, multiply Row 1 of the matrix by Column 1 of the vector.

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 6 \end{bmatrix} = 1 * 5 + 2 * 6 = 17$$

Next, multiply Row 2 of the matrix by Column 1 of the vector.

$$\begin{bmatrix} 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 6 \end{bmatrix} = 3 * 5 + 4 * 6 = 39$$

Finally, write the matrix-vector product.

$$\boldsymbol{Av} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 17 \\ 39 \end{bmatrix}$$

Example II:

Matrix – Vector Multiplication

$$A_{2X3} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 4 \end{bmatrix} \qquad v_{3X1} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \rightarrow \text{“Weights”}$$

What is the expected dimension of the output vector?

$$A_{2X3} v_{3X1} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 4 \end{bmatrix}_{2X3} \cdot \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}_{3X1}$$

$$= \begin{bmatrix} 1*2+2*(-2)+(-1)*1 \\ 2*2+0*(-2)+4*1 \end{bmatrix}_{2X1}$$

$$= \begin{bmatrix} -3 \\ 8 \end{bmatrix}_{2X1}$$

Example II:

Matrix – Vector Multiplication

$$A_{2X3} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 4 \end{bmatrix} \qquad v_{3X1} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

"Weights"

What is the expected dimension of the output vector?

$$A_{2X3}v_{3X1} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 4 \end{bmatrix}_{2X3} \cdot \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}_{3X1}$$

$$= \begin{bmatrix} 1 * 2 + 2 * (-2) + (-1) * 1 \\ 2 * 2 + 0 * (-2) + 4 * 1 \end{bmatrix}_{2X1}$$

$$= 2 * \begin{bmatrix} 1 \\ 2 \end{bmatrix}_{2X1} + (-2) * \begin{bmatrix} 2 \\ 0 \end{bmatrix}_{2X1} + 1 * \begin{bmatrix} -1 \\ 4 \end{bmatrix}_{2X1}$$

$$= \begin{bmatrix} -3 \\ 8 \end{bmatrix}_{2X1}$$

BREAK

MATRIX – MATRIX MULTIPLICATION

THE UNIVERSITY OF CHICAGO

Definition:

Matrix – Matrix Multiplication

Multiplication of matrices A and B can be defined only if they are of the proper sizes.

$$Y_{mxn} = A_{mxk} B_{kxn}$$

Definition:

Matrix – Matrix Multiplication

The product element in row i and column j (i.e., a_{ij}) is the sum of the products of corresponding elements from row i of $\boldsymbol{A}$ and column j of $\boldsymbol{B}$

Example for multiplying matrix $\boldsymbol{A}$ which contains 2 rows and matrix $\boldsymbol{B}$ which contains 3 columns

$$\begin{bmatrix} Row1 \\ Row2 \end{bmatrix} \begin{bmatrix} Col1 & Col2 & Col3 \end{bmatrix}$$
$$= \begin{bmatrix} Row1 \cdot Col1 & Row1 \cdot Col2 & Row1 \cdot Col3 \\ Row2 \cdot Col1 & Row2 \cdot Col2 & Row2 \cdot Col3 \end{bmatrix}$$

Example: Matrix – Matrix Multiplication

The product element in row i and column j (i.e., a_{ij}) is the sum of the products of corresponding elements from row i of $\boldsymbol{A}$ and column j of $\boldsymbol{B}$

$$A_{2X2} = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} \qquad B_{2X2} = \begin{bmatrix} -1 & 0 \\ 5 & 1 \end{bmatrix}$$

$$A_{2X2}B_{2X2} = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 5 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 9 & 2 \\ 11 & 3 \end{bmatrix}$$

Row 2 → $i = 2$
Column 1 → $j = 1$

$$4 * -1 + 3 * 5 = 11$$

Example: Matrix – Matrix Multiplication

$$A_{2X3} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad B_{3X2} = \begin{bmatrix} 2 & 4 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}$$

What is the expected shape?

$$\begin{bmatrix} Row1 \\ Row2 \end{bmatrix} \begin{bmatrix} Col1 & Col2 \end{bmatrix} = \begin{bmatrix} Row1 \cdot Col1 & Row1 \cdot Col2 \\ Row2 \cdot Col1 & Row2 \cdot Col2 \end{bmatrix}$$

$$A_{2X3}B_{3X2}$$

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

$$= \begin{bmatrix} 1*2+2*1+3*0 & 1*4+2*(-1)+3*0 \\ 4*2+5*1+6*0 & 4*4+5*(-1)+6*0 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 2 \\ 13 & 11 \end{bmatrix}$$

Example: Matrix – Matrix Multiplication

$$A_{2X3} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad B_{3X2} = \begin{bmatrix} 2 & 4 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}$$

What is the expected shape?

$$\begin{bmatrix} Row1 \\ Row2 \end{bmatrix} \begin{bmatrix} Col1 & Col2 \end{bmatrix} = \begin{bmatrix} Row1\ X\ Col1 & Row1\ X\ Col2 \\ Row2\ X\ Col1 & Row2\ X\ Col2 \end{bmatrix}$$

$$A_{2X3}B_{3X2} = \begin{bmatrix} 1*2+2*1+3*0 & 1*4+2*(-1)+3*0 \\ 4*2+5*1+6*0 & 4*4+5*(-1)+6*0 \end{bmatrix}$$

$$= \begin{bmatrix} 2*\begin{bmatrix} 1 \\ 4 \end{bmatrix}_{2X1} + 1*\begin{bmatrix} 2 \\ 5 \end{bmatrix}_{2X1} + 0*\begin{bmatrix} 3 \\ 6 \end{bmatrix}_{2X1} & 4*\begin{bmatrix} 1 \\ 4 \end{bmatrix}_{2X1} + (-1)*\begin{bmatrix} 2 \\ 5 \end{bmatrix}_{2X1} + 0*\begin{bmatrix} 3 \\ 6 \end{bmatrix}_{2X1} \end{bmatrix}$$

Properties of Matrices –
Matrix Multiplication

Given

$$A_{2X2} = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$$

$$B_{2X2} = \begin{bmatrix} -1 & 0 \\ 5 & 1 \end{bmatrix}$$

$$A_{2X2}B_{2X2} = \begin{bmatrix} 9 & 2 \\ 11 & 35 \end{bmatrix}$$

Check

$$B_{2X2}A_{2X2} = \begin{bmatrix} -1 & 0 \\ 5 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -2 \\ 9 & 13 \end{bmatrix}$$

Check

$$B_{MXN}A_{NXM} \neq A_{NXM}B_{MXN}$$

Not Commutative

Properties of Matrices –
Matrix Multiplication

If $\boldsymbol{AB} = \boldsymbol{AC}$ then it is **not true** that $\boldsymbol{B} = \boldsymbol{C}$

If $\boldsymbol{AB} = \boldsymbol{0}$, we **cannot** conclude that either
$\boldsymbol{A} = \boldsymbol{0}\ or\ \boldsymbol{B} = \boldsymbol{0}$

$$A_{2x2}B_{2x2} = \begin{bmatrix} -1 & 4 \\ 3 & -12 \end{bmatrix}\begin{bmatrix} 4 & 8 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$A_{2x2}C_{2x2} = \begin{bmatrix} -1 & 4 \\ 3 & -12 \end{bmatrix}\begin{bmatrix} 12 & 16 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Check

Definition:

Transpose

The transpose A^T of a matrix A is an operation in which the terms above and below the diagonal are interchanged.

Example I

$$A_{2X3} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 2 \end{bmatrix} \qquad A^T_{3X2} = \begin{bmatrix} 1 & 3 \\ 2 & 3 \\ 3 & 2 \end{bmatrix}$$

Example II

$$A_{2X2} = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} \qquad A^T_{2X2} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$$

Note:

For any matrix A_{NXM}: $A^T A$ and AA^T are square matrices of size M and N, respectively.

Definition:

Trace

The **Trace** of a square matrix A_{nXn} is the sum of its diagonal elements.

$$tr\ A = \sum_{i=1}^{N} a_{ii}$$

Example

$$\text{tr}\begin{bmatrix} 2 & 7 & 6 \\ 9 & 5 & 1 \\ 3 & 2 & 8 \end{bmatrix} = 2 + 5 + 8 = 15$$

Definition:

Diagonal Matrix

A diagonal matrix has nonzero terms only along its main diagonal.

The matrix A_{NXN} is diagonal if $a_{ij} = 0 \; if \; i \neq j$

The sum and product of diagonal matrices are also diagonal.

Example $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}$

Examples:

Diagonal Matrix

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$A + B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Recall:
The sum and product of diagonal matrices are also diagonal.

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

Definition:

Identity Matrix

The *identity* matrix I is a square diagonal matrix with 1 on the main diagonal.

Example

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{A \cdot I = I \cdot A = A}$$

Definition:

Symmetric Matrix

A symmetric matrix is one for which $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

$$\mathbf{A}^{\mathrm{T}} = \mathbf{A}$$

$$or$$

$$a_{ij} = a_{ji}$$

Example

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 2 & 5 \\ 4 & 5 & -1 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 2 & 5 \\ 4 & 5 & -1 \end{bmatrix}$$

Definition:

Triangular Matrices

A **lower triangular** matrix is one in which all entries above the main diagonal are zero. Lower triangular matrices are often denoted by L.

$$L = \begin{bmatrix} \ell_{1,1} & & & & 0 \\ \ell_{2,1} & \ell_{2,2} & & & \\ \ell_{3,1} & \ell_{3,2} & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{n,1} & \ell_{n,2} & \ldots & \ell_{n,n-1} & \ell_{n,n} \end{bmatrix}$$

The transpose of a **lower triangular** matrix is an **upper triangular** matrix (Check)

Example $\begin{bmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 4 & 5 & 0 \end{bmatrix}$

Definition:

Triangular Matrices

An **upper triangular** matrix is one in which all entries below the main diagonal are zero. Upper triangular matrices are often denoted by U.

$$U = \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & \cdots & u_{1,n} \\ & u_{2,2} & u_{2,3} & \cdots & u_{2,n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & u_{n-1,n} \\ 0 & & & & u_{n,n} \end{bmatrix}$$

The transpose of an **upper triangular** matrix is **lower triangular** matrix (Check)

Example

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

Definition:

Orthogonal Matrices

A matrix is orthogonal if all its columns are <u>orthogonal</u> unit vectors.

$$\boldsymbol{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{bmatrix}$$

Examples

$$\mathbf{Q} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{Q} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad \mathbf{Q} = \begin{bmatrix} 3/7 & 2/7 & 6/7 \\ -6/7 & 3/7 & 2/7 \\ 2/7 & 6/7 & -3/7 \end{bmatrix}$$

Check

Definition:

Inverse Matrix

- What is the inverse of 3 ?

$1/3$ ➔ $3 * 1/3 = 1$

- What should the inverse of matrix $\boldsymbol{A}$ do?

$$\boldsymbol{I} = \boldsymbol{A}\ (Inverse\ Matrix)$$

Definition:

Matrix A has an inverse $\boldsymbol{A}^{-1}$ if

$$\boldsymbol{I} = \mathbf{A}\,\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A}$$

Examples:

Inverse Matrix

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

$$\underset{\boldsymbol{A}}{\begin{bmatrix}2 & 2\\-2 & 2\end{bmatrix}}\underset{\boldsymbol{A^{-1}}}{\begin{bmatrix}\frac{1}{4} & -\frac{1}{4}\\\frac{1}{4} & \frac{1}{4}\end{bmatrix}}=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}=\boldsymbol{I}$$

Definitions:

Invertible and Singular Matrices

- If A^{-1} exists ➔ A is **Invertible**
- If A^{-1} doesn't exist ➔ A is **Singular**

- **Independent** column vectors ➔ A is **Invertible**
- **Dependent** column vectors ➔ A is **Singular**

Cosine Similarity

THE UNIVERSITY OF CHICAGO

Real World Application - Measuring Distance and Similarity

- We saw that the Norm can help measure the **distance** between two vectors.
- The Cosine of the angle between two vectors can be used to measure the **similarity** between the two vectors.
  - **Similar:** If vectors v and u are close, $\boldsymbol{\theta} \sim 0$ and $\text{Cos}\,\boldsymbol{\theta} \sim 1$
  - **Dissimilar:** If vectors v and u are orthogonal $\boldsymbol{\theta} = 90^{\circ}$ and $\text{Cos}\,\boldsymbol{\theta} = 0$

$$0 \leq |\cos \boldsymbol{\theta}| \leq 1$$

Measuring Distance and Similarity – Text Mining Example

Word – Document Matrix

		Doc 1	Doc 2	Doc 3
e.g., Ball	Word 1	8	0	2
e.g., Yard	Word 2	8	0	2
e.g., Math	Word 3	0	2	0
		e.g., ESPN	**e.g., Nature**	**e.g., NYT Sports**

- The documents are represented by a vector in $\mathbb{R}^3$
- Realistically, we may have thousands of documents and words.

Measuring Distance and Similarity – Text Mining Example

Word – Document Matrix

		Doc 1	Doc 2	Doc 3
e.g., Ball	Word 1	8	0	2
e.g., Yard	Word 2	8	0	2
e.g., Math	Word 3	0	2	0
		e.g., ESPN	**e.g., Nature**	**e.g., NYT Sports**

$\|\boldsymbol{Doc\ 1}\| = \mathbf{11.3}$ $\|\boldsymbol{Doc\ 2}\| = \mathbf{2}$ $\|\boldsymbol{Doc\ 3}\| = \mathbf{2.8}$

$\|\boldsymbol{Doc\ 1 - Doc3}\| = \mathbf{8.4}$ $\|\boldsymbol{Doc\ 2 - Doc3}\| = \mathbf{3.4}$

Using the Euclidean distance Documents 1 and 3 look **dissimilar**, and Documents 2 and 3 look **similar**. This is just due to the length of the documents!

Measuring Distance and Similarity – Text Mining Example

Word – Document Matrix

		Doc 1	Doc 2	Doc 3
e.g., Ball	Word 1	8	0	2
e.g., Yard	Word 2	8	0	2
e.g., Math	Word 3	0	2	0
		e.g., ESPN	**e.g., Nature**	**e.g., NYT Sports**

$\text{Cos}\,\boldsymbol{\theta}_{12} = \mathbf{0}$ $\text{Cos}\,\boldsymbol{\theta}_{13} = \mathbf{1}$ $\text{Cos}\,\boldsymbol{\theta}_{23} = \mathbf{0}$

Using the <u>cosine of the angle</u> between document vectors Documents 1 and 3 are **similar** to each other and **dissimilar** to Document 2.

PYTHON EXAMPLE

*.linalg

Linear algebra (scipy.linalg)

Linear algebra functions.

See also:
numpy.linalg for more linear algebra functions. Note that although scipy.linalg imports most of them, identically named func
from scipy.linalg may offer more or slightly differing functionality.

Basics

inv(a[, overwrite_a, check_finite])	Compute the inverse of a matrix.
solve(a, b[, sym_pos, lower, overwrite_a, ...])	Solve the equation `a x = b` for `x`.
solve_banded(l_and_u, ab, b[, overwrite_ab, ...])	Solve the equation a x = b for x, assuming a is banded matrix.
solveh_banded(ab, b[, overwrite_ab, ...])	Solve equation a x = b.
solve_circulant(c, b[, singular, tol, ...])	Solve C x = b for x, where C is a circulant matrix.
solve_triangular(a, b[, trans, lower, ...])	Solve the equation *a x* = *b* for *x*, assuming a is a triangular matrix.
solve_toeplitz(c_or_cr, b[, check_finite])	Solve a Toeplitz system using Levinson Recursion
det(a[, overwrite_a, check_finite])	Compute the determinant of a matrix
norm(a[, ord, axis, keepdims])	Matrix or vector norm.
lstsq(a, b[, cond, overwrite_a, ...])	Compute least-squares solution to equation Ax = b.
pinv(a[, cond, rcond, return_rank, check_finite])	Compute the (Moore-Penrose) pseudo-inverse of a matrix.

Linear algebra (numpy.linalg)¶

Matrix and vector products

dot(a, b[, out])	Dot product of two arrays.
vdot(a, b)	Return the dot product of two vectors.
inner(a, b)	Inner product of two arrays.
outer(a, b[, out])	Compute the outer product of two vectors.
matmul(a, b[, out])	Matrix product of two arrays.
tensordot(a, b[, axes])	Compute tensor dot product along specified axes for arrays >= 1-D.
einsum(subscripts, *operands[, out, dtype, ...])	Evaluates the Einstein summation convention on the operands.
linalg.matrix_power(M, n)	Raise a square matrix to the (integer) power *n*.
kron(a, b)	Kronecker product of two arrays.

Decompositions

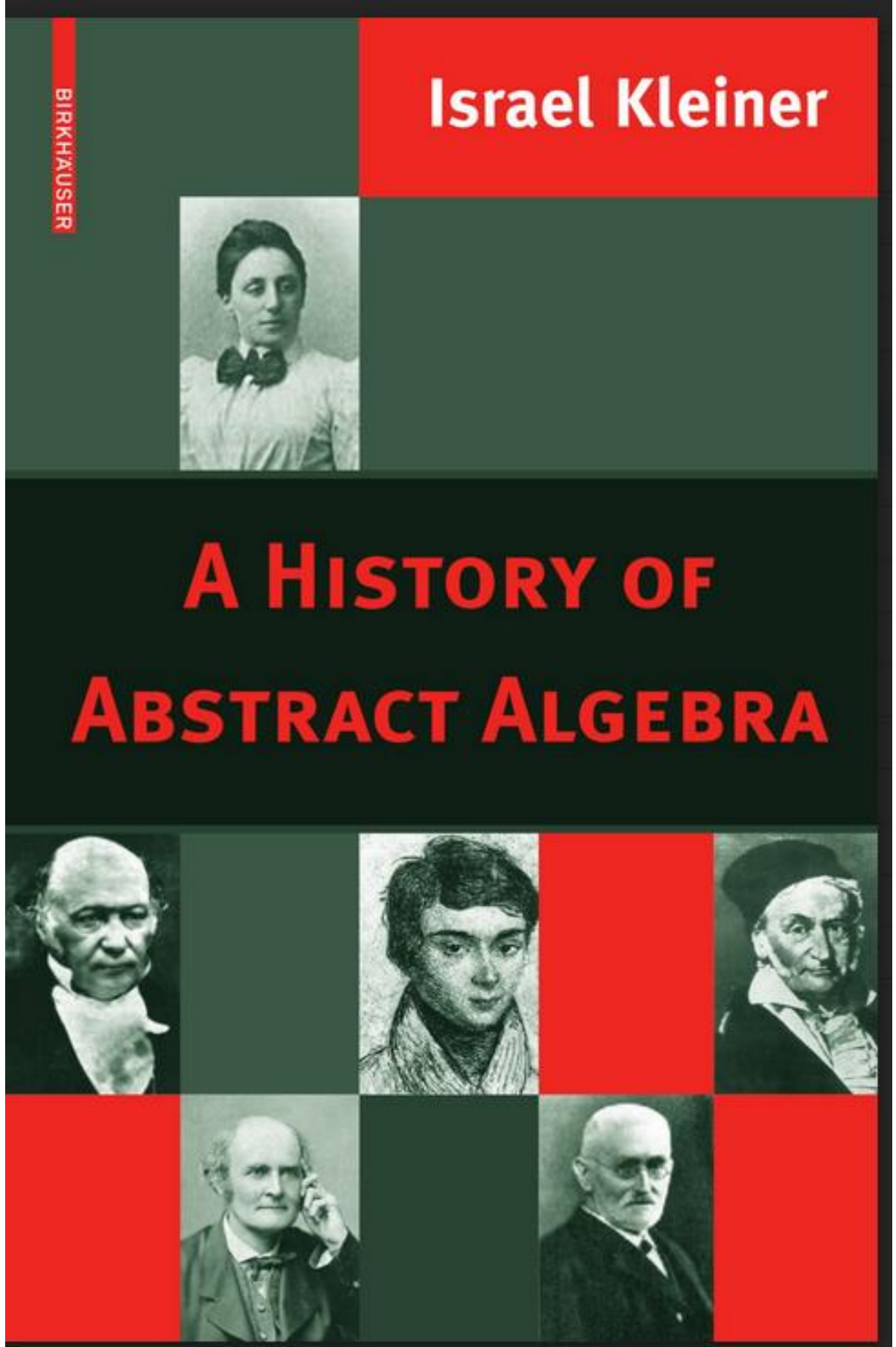
linalg.cholesky(a)	Cholesky decomposition.
linalg.qr(a[, mode])	Compute the qr factorization of a matrix.
linalg.svd(a[, full_matrices, compute_uv])	Singular Value Decomposition.

A Brief History of Linear Algebra

THE UNIVERSITY OF CHICAGO

Linear Algebra – Brief History

- **Around 4000 years ago,** the people of Babylon knew how to solve a simple 2X2 system of linear equations with two unknowns.
- **Around 200 BC,** the Chinese published that "Nine Chapters of the Mathematical Art," they displayed the ability to solve a 3X3 system of equations
- The power and progress in linear algebra did not come to fruition until the **late 17th** century.
  - Leibnitz - Determinants
  - Lagrange - Lagrange multipliers,
  - Cramer – Cramer's Law
  - Euler - System of equations doesn't necessarily have to have a solution
- **19th century**
  - Gauss introduced a procedure to be used for solving a system of linear equation (Gaussian elimination)
- **1848** J.J. Sylvester introduced the term "matrix," the Latin word for womb, as a name for an array of numbers.
- **1855** Introduced Arthur Cayley Matrix multiplication or matrix algebra, Identity matrix and matrix inverse
- **Post WWII** - With the advancement of technology using the methods of Cayley, Gauss, Leibnitz, Euler, and other determinants and linear algebra moved forward more quickly and more effective.


Israel Kleiner
BIRKHÄUSER
A HISTORY OF
ABSTRACT ALGEBRA

EXTRA SLIDES

THE UNIVERSITY OF CHICAGO