NUS | Computing
National University of Singapore

Linear Algebra

Introduction to Computing Foundations

Sirigina Rajendra Prasad

11-January-2024

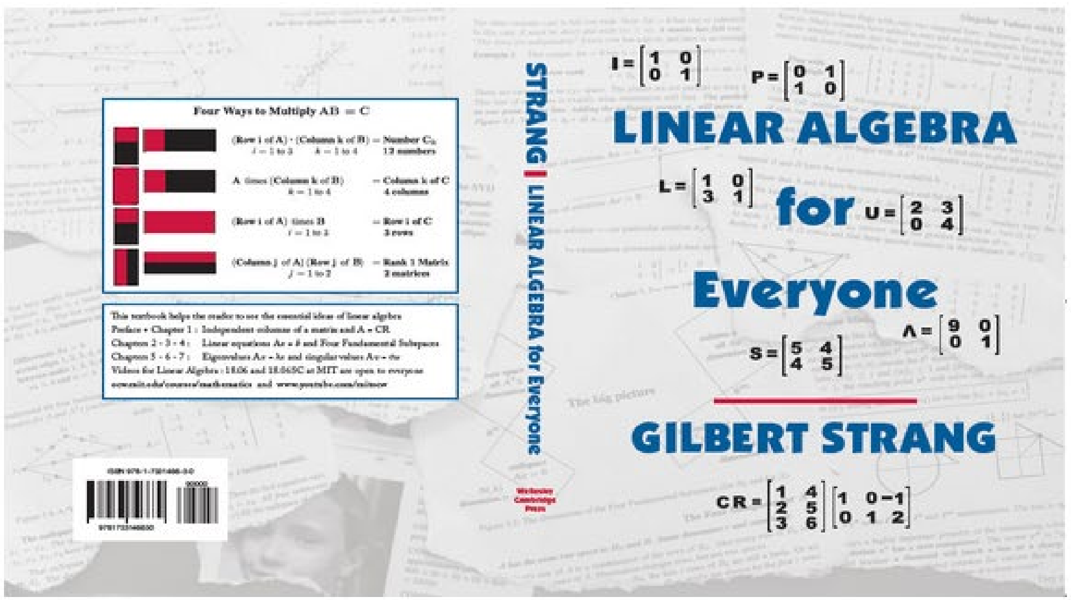
Textbook

https://math.mit.edu/~gs/everyone/

Linear Algebra for Everyone [Gilbert Strang]

ISBN 978-1-7331466-3-0 **September 2020**

[Wellesley-Cambridge Press] [gilstrang@gmail.com]



https://ocw.mit.edu/courses/18-06-linear-algebra-spring-2010/video_galleries/video-lectures/

Textbook

https://textbooks.math.gatech.edu/ila/

Interactive Linear Algebra
Dan Margalit, Joseph Rabinoff

PDF version

≡ Index ‹ Prev ^ Up Next ›

Interactive Linear Algebra

Dan Margalit
School of Mathematics
Georgia Institute of Technology

Joseph Rabinoff
School of Mathematics
Georgia Institute of Technology

June 3, 2019

References:

1. https://github.com/kenjihiranabe/The-Art-of-Linear-Algebra

2. https://math.mit.edu/~gs/linearalgebra/ila5/indexila5.html

3. Lecture Materials by Dr. Victor Tan et al. (MA1101R) Linear Algebra 1, NUS

Agenda

- Part I:
  - ➢ Motivation
  - ➢ Vectors

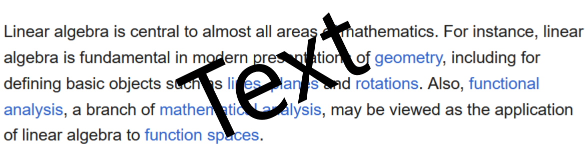
- Part II:
  - ➢ Matrices

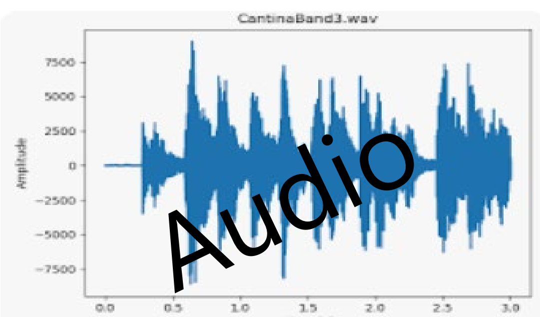
- Part III:
  - ➢ Projection
  - ➢ Linear Systems
  - ➢ Matrix Decompositions

What is linear algebra?

Why Linear Algebra?

Raw Data


Linear algebra is central to almost all areas of mathematics. For instance, linear algebra is fundamental in modern presentations of geometry, including for defining basic objects such as lines, planes and rotations. Also, functional analysis, a branch of mathematical analysis, may be viewed as the application of linear algebra to function spaces.



→ Algorithms →

Feature Matrix

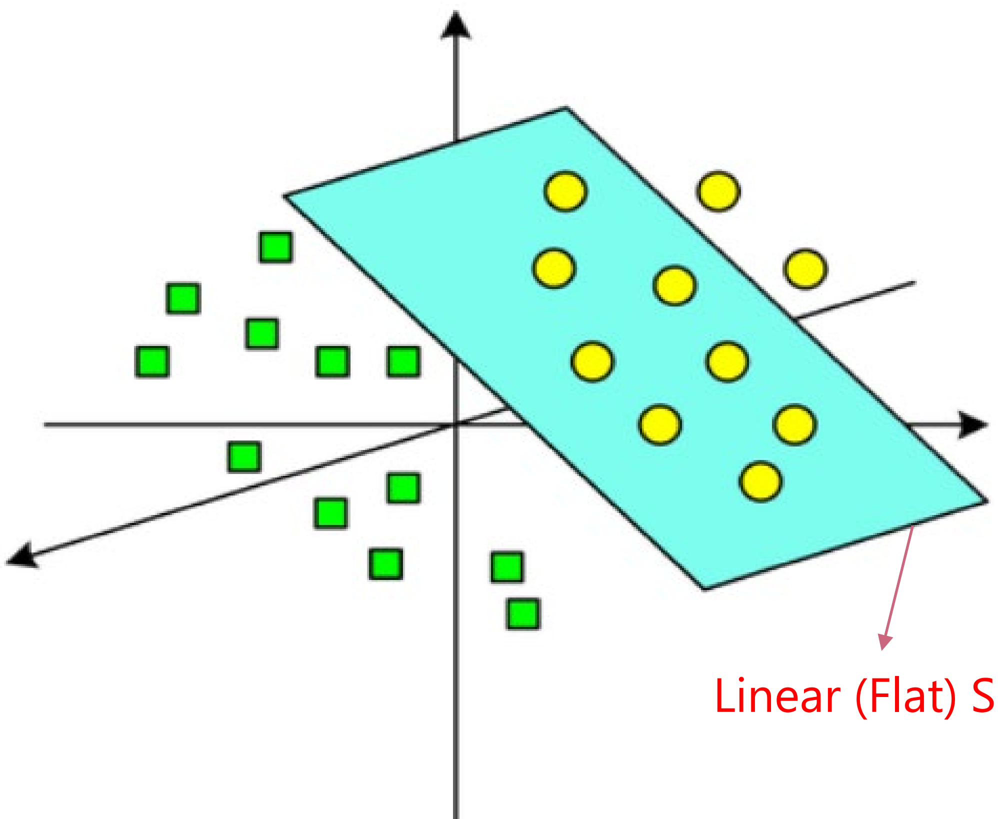
← Feature Vector →
← Feature Vector →
← Feature Vector →
← Feature Vector →
← Feature Vector →
← Feature Vector →
← Feature Vector →

Each vector corresponds to an image/document/audio

Why Linear Algebra?

"Learning" flat surfaces is easy

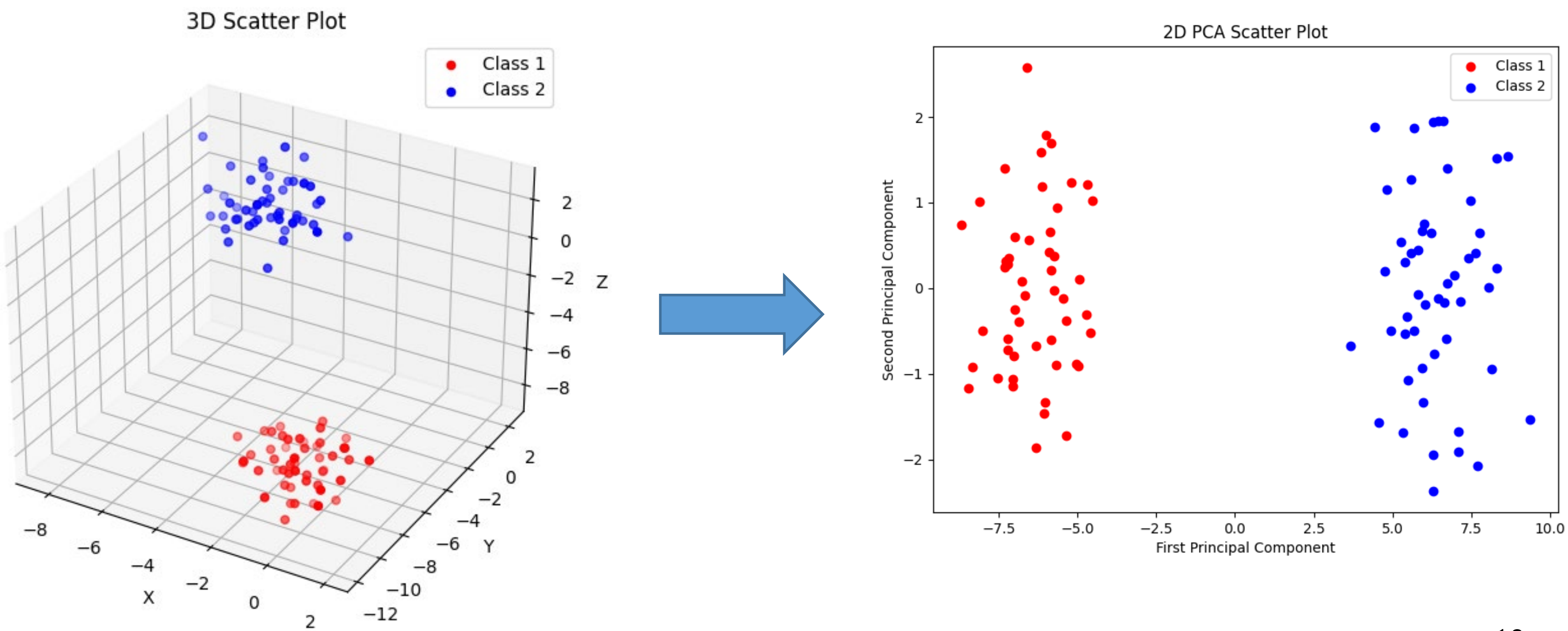
- both for humans and machines



Each point represents a vector

Why Linear Algebra?

Projection of data to lower dimensions



Lower number of dimensions:

Good for Visualization

Efficient Reasoning

Linear Algebra: Key Tool(s)

Matrix Factorization

Lowers computational complexity
Numerical stability, etc.

Matrix Factorization

Decomposing a matrix into product of multiple matrices

(Orthogonal Matrix)(Triangular Matrix)

(Orthogonal Matrix) (Diagonal Matrix) (Orthogonal Matrix)

Milestones:

Vectors
- Norm
- Independent Vectors
- Dimension
- Span
- Space

Matrices
- Types of Matrices
- Matrix Addition
- Matrix Multiplication
- Determinant
- Eigen Values
- Eigen Vectors
- Rank

Projection onto a space
Least Squares Solution
QR decomposition
***ML Use Cases: Linear Regression, etc.**

Eigen Value Decomposition
Singular Value Decomposition
***ML Use Cases: PCA, etc.**

*We are not doing ML use cases in this bootcamp

Part I:Vectors

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix} \in \mathbb{R}^{\boldsymbol{n}}$$

Vectors: Agenda

- Norm
- Inner and Outer product
- Distance between Vectors
- Vector Spaces
    - ➢ Independence/Dependence
    - ➢ Basis
    - ➢ Span
    - ➢ Dimension

Vectors

	Column Vector	Row Vector
2-Dimensional Vector	$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^{\mathbf{2}}$	$\boldsymbol{x} = [x_1 \quad x_2] \in \mathbb{R}^{\mathbf{2}}$
3-Dimensional Vector	$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^{\mathbf{3}}$	$\boldsymbol{x} = [x_1 \quad x_2 \quad x_3] \in \mathbb{R}^{\mathbf{3}}$
n-Dimensional Vector	$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \in \mathbb{R}^{\boldsymbol{n}}$	$\boldsymbol{x} = [x_1 \quad x_2 \ \dots \quad x_n] \in \mathbb{R}^{\boldsymbol{n}}$

Unless explicitly stated, a vector is a column vector

Transpose of a Vector

- Transpose of a vector:

$$\begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$$

$$\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

$\|\boldsymbol{x}\|_p$: p-norm of vector $\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}$

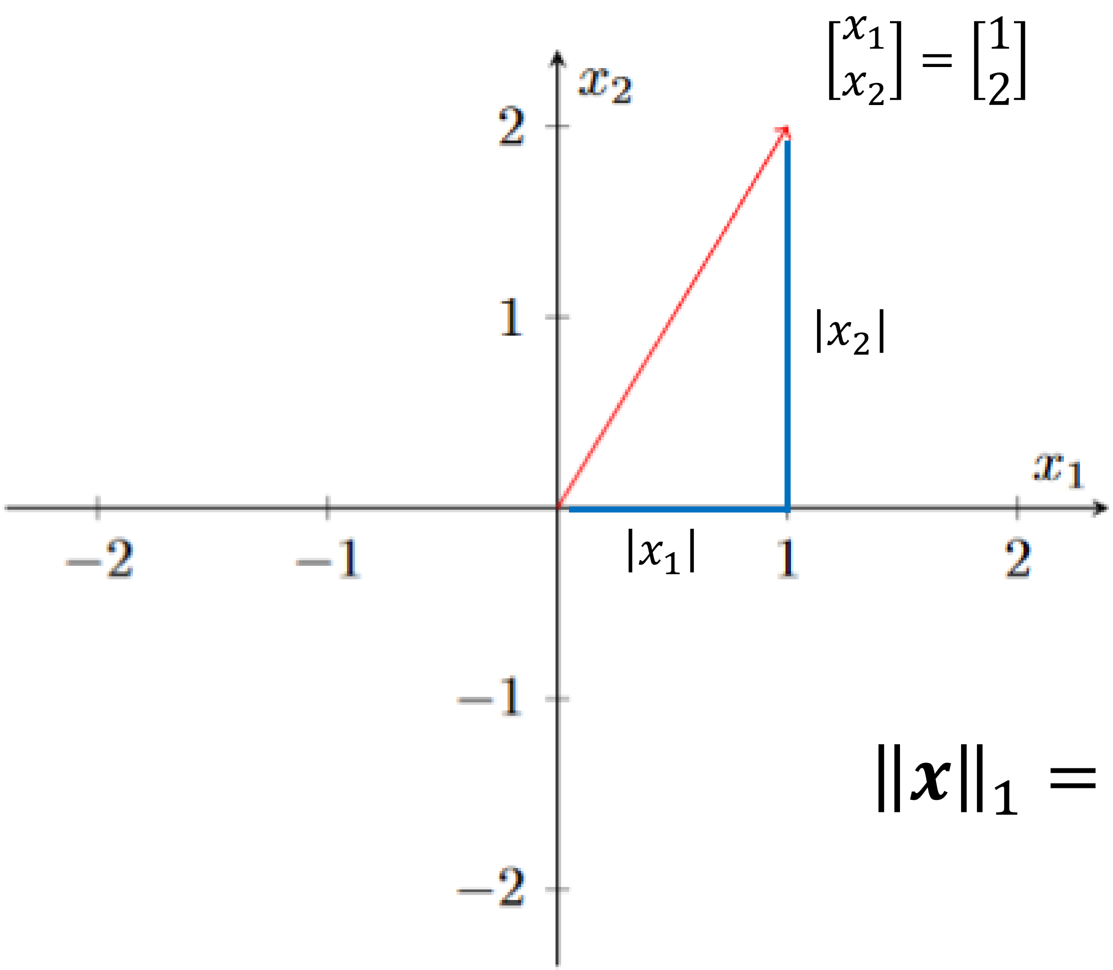
$$\|\boldsymbol{x}\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots |x_n|^p}$$

$$= \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}$$

p-norm of a vector is a scalar

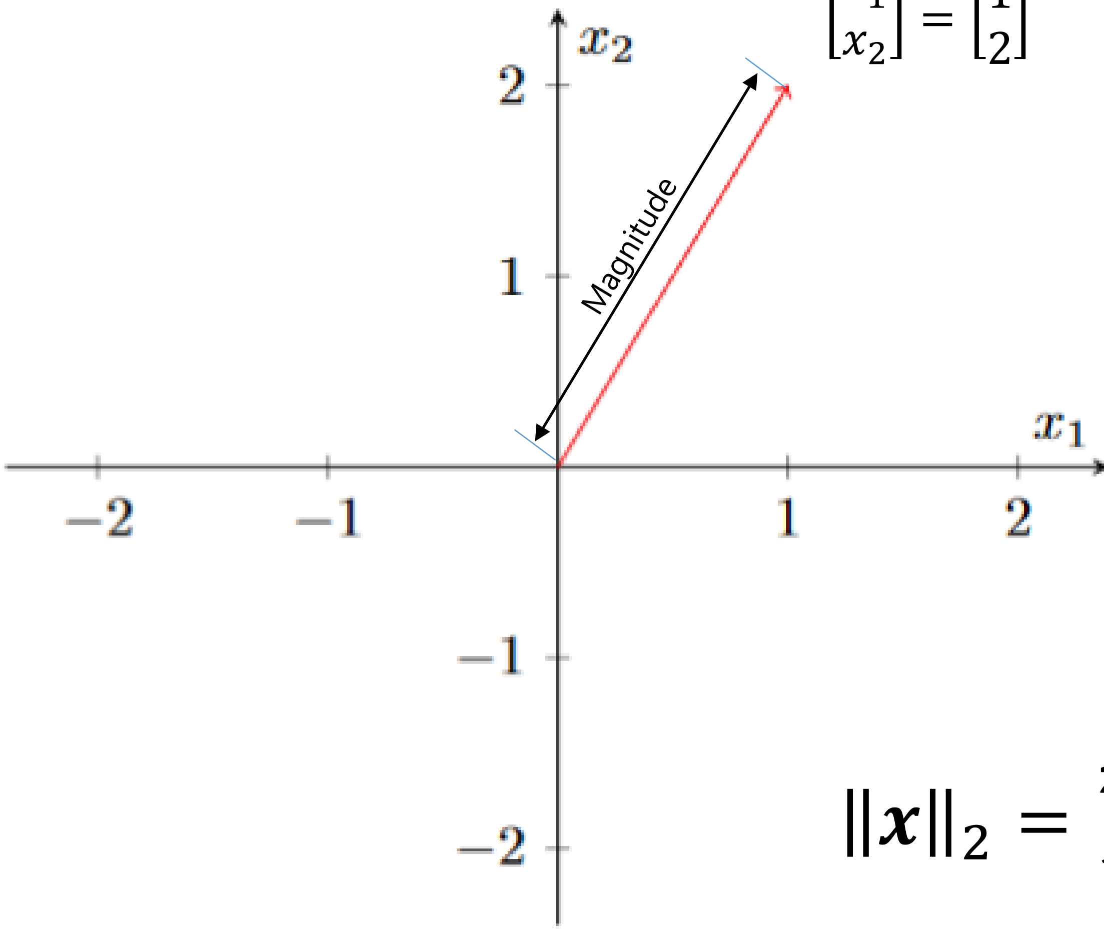
p	**Norm**
1	Manhattan Norm (Taxi-Cab Norm)
2	Euclidean Norm
∞	Infinity Norm or Max norm

Manhattan Norm ($\|\boldsymbol{x}\|_1$) in Two Dimensions



$$\|\boldsymbol{x}\|_1 = |x_1| + |x_2| = 3$$

Euclidean Norm ($\|\boldsymbol{x}\|_2$) in Two Dimensions

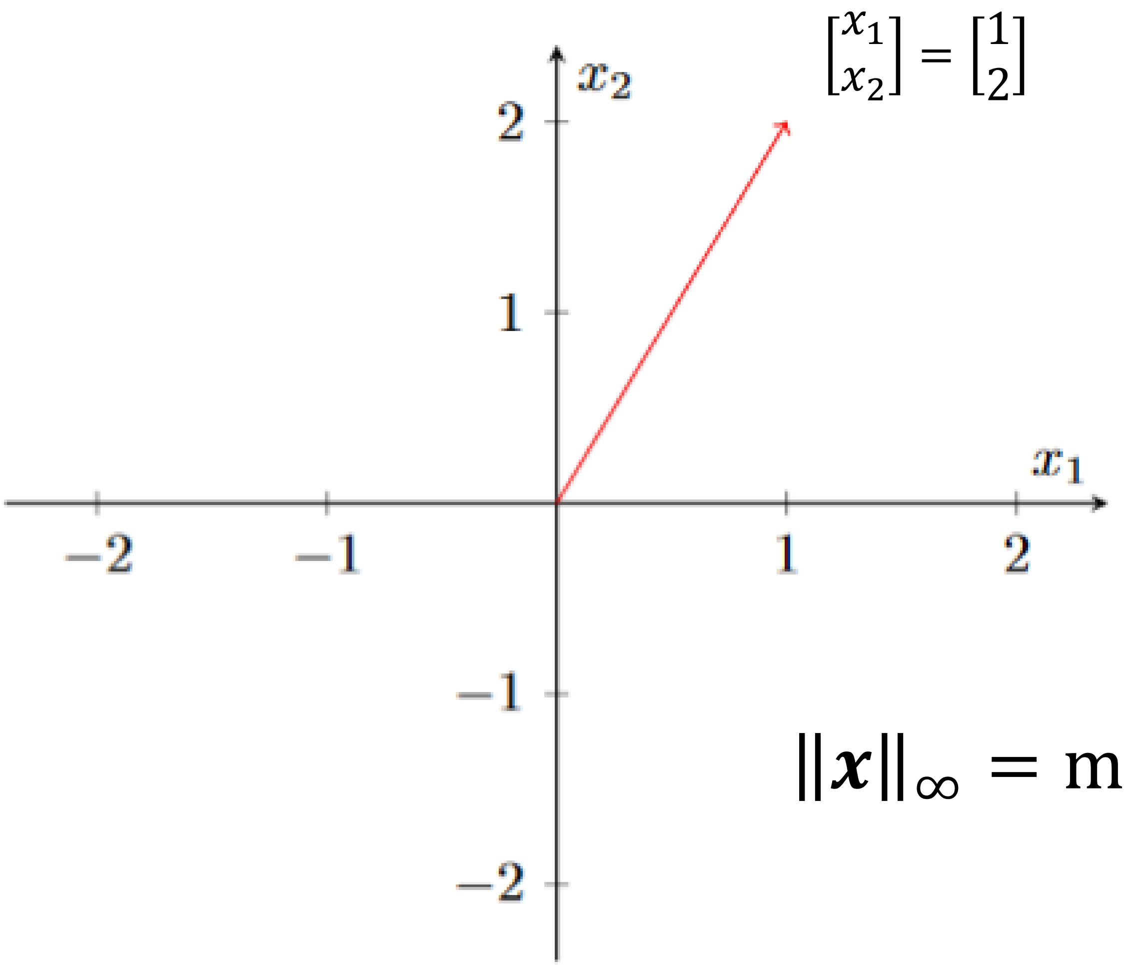


$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\|\boldsymbol{x}\|_2 = \sqrt[2]{x_1^2 + x_2^2} = \sqrt[2]{5}$$

Cartesian Coordinates: X and Y- axes

Infinity Norm ($\|\boldsymbol{x}\|_\infty$) in Two Dimensions



$$\|\boldsymbol{x}\|_\infty = \max(|x_1|, |x_2|) = 2$$

Unit Vector

- If $\|\boldsymbol{x}\|_2 = 1$, then $\boldsymbol{x}$ is a unit vector

- Euclidean norm (or magnitude) of unit vector is 1

- Identify the unit vectors?

  A. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

  B. $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

  C. $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

  D. $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

How to convert a vector to unit vector?

- Divide each element of the vector with its Euclidean norm

If $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$, then $\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}$ is a unit vector

Ex: $\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\|\boldsymbol{x}\|_2 = \sqrt{2}$ $\Rightarrow$ $\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ is a unit vector

Operations on Vectors

- Addition
- Subtraction
- Product
- Distance between vectors

$$\boldsymbol{x}_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ . \\ . \\ . \\ x_{n1} \end{bmatrix} \text{ and } \boldsymbol{x}_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ . \\ . \\ . \\ x_{n2} \end{bmatrix}$$

Addition and Subtraction

- **Addition**:

  ➢ $\boldsymbol{x}_1 + \boldsymbol{x}_2 = \begin{bmatrix} x_{11} + x_{12} \\ x_{21} + x_{22} \\ \cdot \\ \cdot \\ \cdot \\ x_{n1} + x_{n2} \end{bmatrix}$

- **Subtraction:**

  ➢ $\boldsymbol{x}_1 - \boldsymbol{x}_2 = \begin{bmatrix} x_{11} - x_{12} \\ x_{21} - x_{22} \\ \cdot \\ \cdot \\ \cdot \\ x_{n1} - x_{n2} \end{bmatrix}$

Two vectors can be added or subtracted only if they are of same size

Product

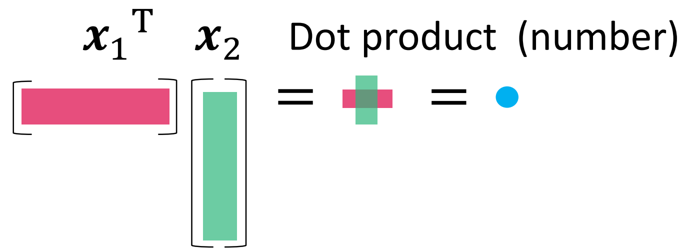
- Product of a scalar and a vector
    - If c is a scalar, then
        - $c.\boldsymbol{x}_1 = \begin{bmatrix} c.x_{11} \\ c.x_{21} \\ . \\ . \\ . \\ c.x_{n1} \end{bmatrix}$
- Product of two vectors
    - Inner Product or Dot Product
    - Outer Product or Cross Product

Inner Product or Dot Product of Vectors

- Inner/Dot product is defined as: $\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = \boldsymbol{x}_1{}^T\, \boldsymbol{x}_2 = \boldsymbol{x}_2{}^T\, \boldsymbol{x}_1 = < \boldsymbol{x}_1, \boldsymbol{x}_2 >$

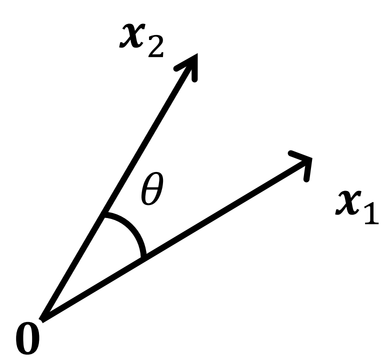
➢ $\boldsymbol{x}_1{}^T\, \boldsymbol{x}_2 = [x_{11} \quad x_{21} \;\cdots\; x_{n1}]\begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix}$

$= \sum_{i=1}^{n} x_{i1} x_{i2}$



➢ $\boldsymbol{x}_2{}^T\, \boldsymbol{x}_1 = [x_{12} \quad x_{22} \;\cdots\; x_{n2}]\begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix}$

$= \sum_{i=1}^{n} x_{i2} x_{i1}$



➢ $< \boldsymbol{x}_1, \boldsymbol{x}_2 > = \|\boldsymbol{x}_1\|_2 \|\boldsymbol{x}_2\|_2 \cos\theta$

Inner Product: Examples

Example 1:

$$\begin{bmatrix}1\\2\\3\end{bmatrix}\cdot\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}^{\boldsymbol{T}}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1 & 2 & 3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=x_1+2x_2+3x_3$$

Example 2:

$$\begin{bmatrix}1\\0\\0\end{bmatrix}\cdot\begin{bmatrix}0\\1\\0\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}^{\boldsymbol{T}}\begin{bmatrix}0\\1\\0\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix}=0+0+0=0$$

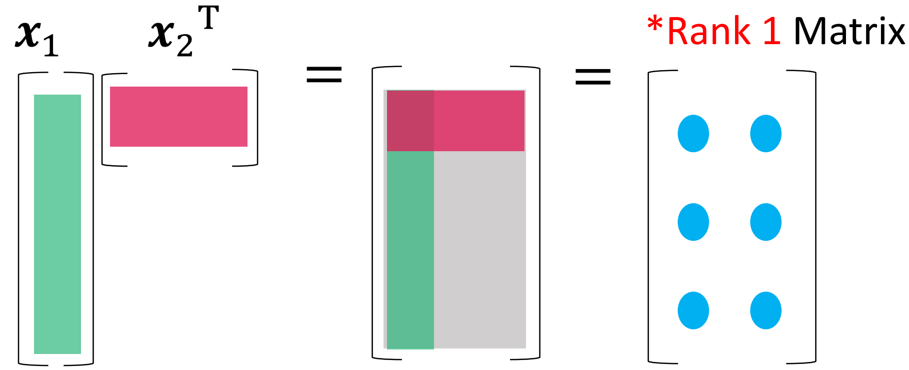
Outer Product of Column Vectors

- Outer Product ($\boldsymbol{x_1} \times \boldsymbol{x_2}$)

  - $\boldsymbol{x_1} \times \boldsymbol{x_2} = \boldsymbol{x_1}\boldsymbol{x_2^T}$

$$= \begin{bmatrix} x_{11} \\ x_{21} \\ \cdot \\ \cdot \\ \cdot \\ x_{n1} \end{bmatrix} [x_{12} \quad x_{22} \;\cdots\; x_{n2}]$$

$$= \begin{bmatrix} x_{11}x_{12} & x_{11}x_{22} & \cdots & x_{11}x_{n2} \\ x_{21}x_{12} & x_{21}x_{22} & \cdots & x_{21}x_{n2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ x_{n1}x_{12} & x_{n1}x_{22} & & x_{n1}x_{n2} \end{bmatrix}$$



* We will discuss Rank 1 matrix later

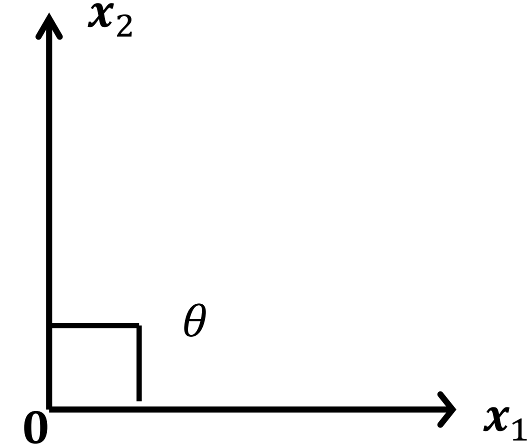
Orthogonal Vectors

- If the inner product of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is zero, then the pair of non-zero vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are orthogonal

$$\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = \boldsymbol{x}_1{}^T \boldsymbol{x}_2$$

$$= \|\boldsymbol{x}_1\|_2 \|\boldsymbol{x}_2\|_2 \cos(90)$$

$$= 0$$



Orthonormal Vectors

- A set of vectors $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n\}$ are called

  orthonormal if

$$\boldsymbol{q}_i . \boldsymbol{q}_j = \begin{cases} 1 & if\ i = j \\ 0 & if\ i \neq j \end{cases}$$

- Identify set of orthonormal vectors

A. $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$

B. $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}-1\\1\end{bmatrix}\right\}$

Orthogonal to Orthonormal Vectors

- Convert each vector to unit vector

- Example:
  - $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}-1\\1\end{bmatrix}\right\}$ are orthogonal vectors. Convert each

  vector in the set to unit vector. The set of unit vectors

  $\left\{\begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}, \begin{bmatrix}-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}\right\}$ are orthonormal

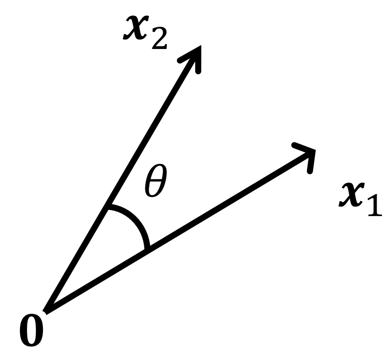
Euclidean Norm and Inner Products

- $\|\boldsymbol{x}\|_2 = \sqrt[2]{|x_1|^2 + |x_2|^2 + \cdots |x_n|^2}$

$$\Rightarrow \|\boldsymbol{x}\|_2^2 = |x_1|^2 + |x_2|^2 + \cdots |x_n|^2$$

$$= \boldsymbol{x}.\boldsymbol{x}$$

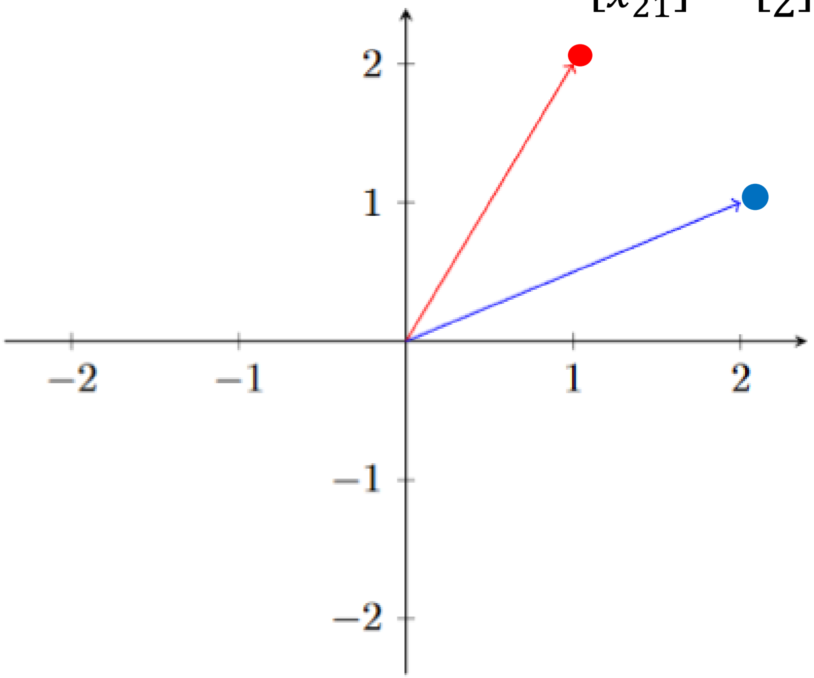
Distance Between Two Vectors



- Manhattan Distance: $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_1$
  - $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_1 = \sum_{i=1}^{n} |x_{i1} - x_{i2}|$
- Euclidean Distance: $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2$
  - $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2 = \sqrt[2]{\sum_{i=1}^{n} (x_{i1} - x_{i2})^2}$
- Angular Distance:
  - $\theta = \frac{1}{\pi} \cos^{-1}\left(\frac{\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle}{\|\boldsymbol{x}_1\|_2 \|\boldsymbol{x}_2\|_2}\right)$

Distance between 2D vectors: Example

$$\begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$



$$\begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

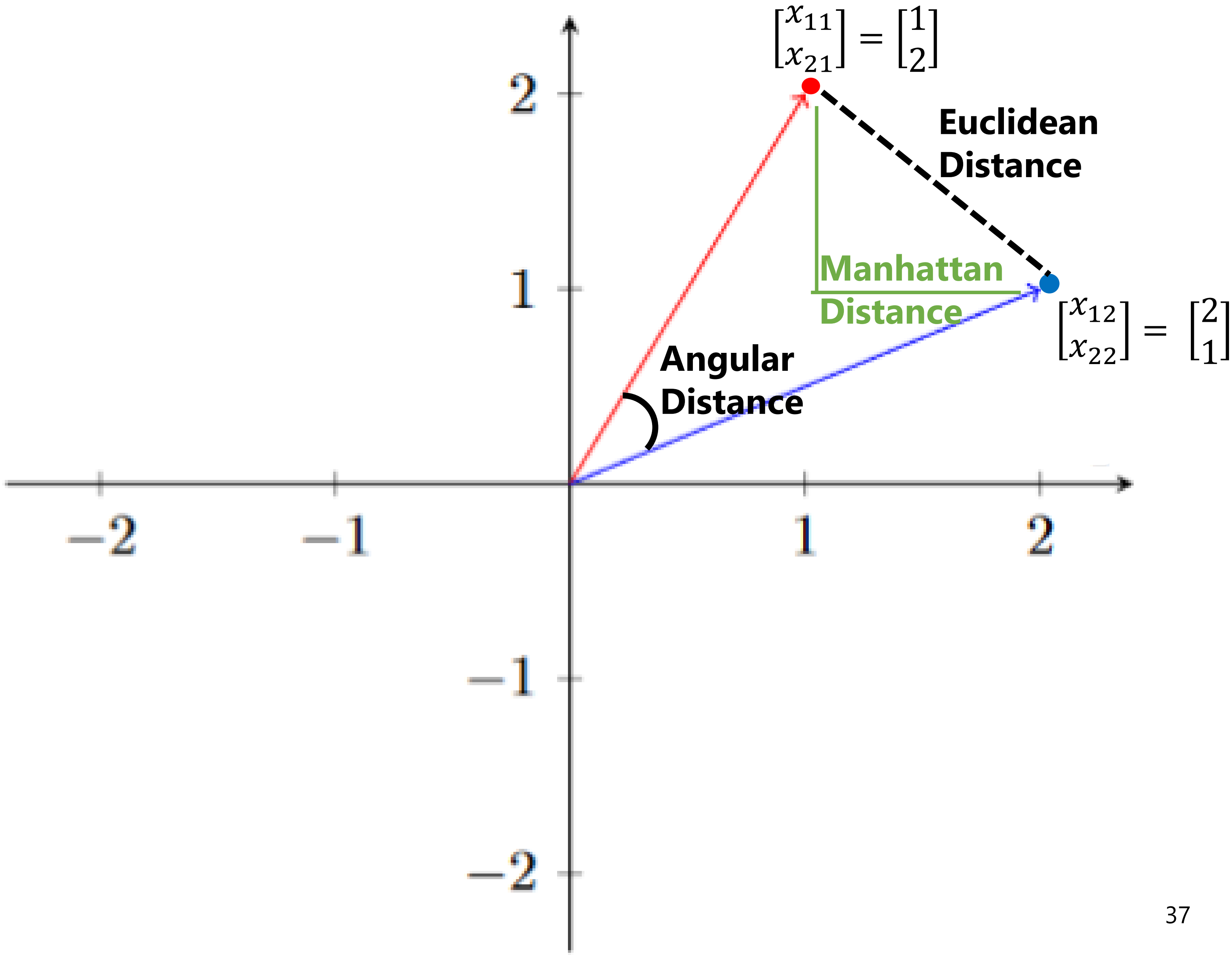
Distance Between 2D Vectors: Example

- Euclidean Distance: $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2$
  - $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2 = \sqrt[2]{(x_{11} - x_{12})^2 + (x_{21} - x_{22})^2}$

- Manhattan Distance: $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_1$
  - $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_1 = |x_{11} - x_{12}| + |x_{21} - x_{22}|$

- Angular Distance:
  - $\theta = \cos^{-1}\left(\frac{\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle}{\|\boldsymbol{x}_1\|_2 \|\boldsymbol{x}_2\|_2}\right)$

Illustration: Distance between 2D vectors

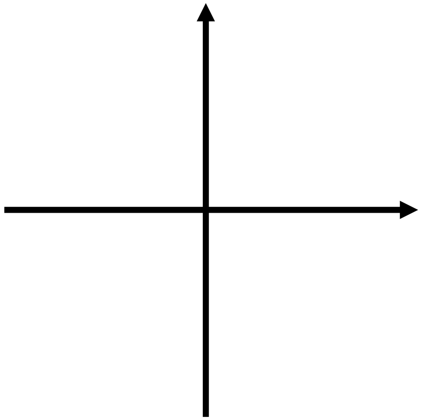


Spaces

Vector Spaces and Subspaces

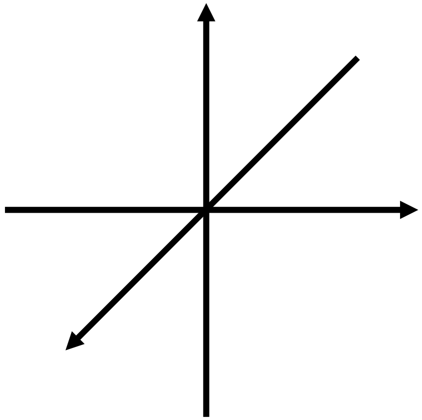
Vector Spaces

2-D Space ($n = 2$)



1-D Space ($n = 1$)

3-D Space ($n = 3$)

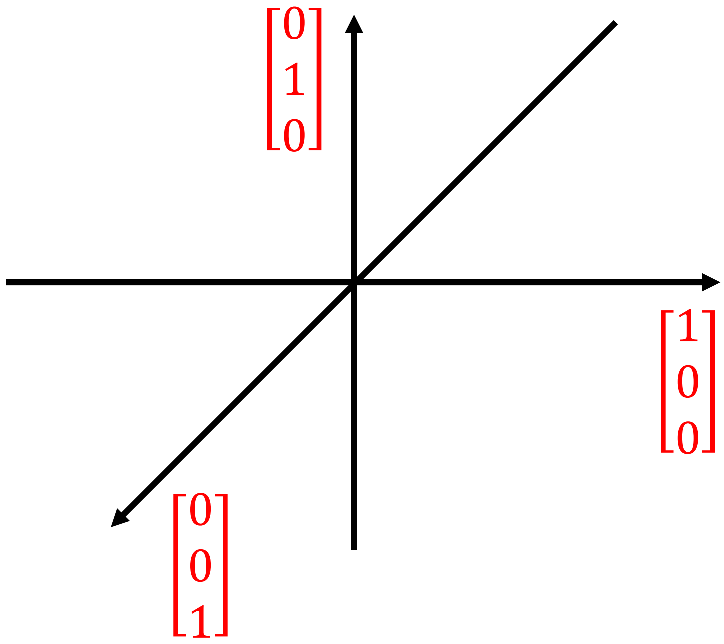


n-D Space for $n > 3$??

Vector Spaces: Vocabulary

Set of Basis Vectors: $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$

3-D Space ($n = 3$)



Basis vectors are **Linearly independent**

Dimensions: 3 (size of set of **basis** vectors)

3-D space is the **Span** of the basis vectors $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$

Euclidean Geometry Vs Linear Algebra: Vocabulary

Euclidean Geometry	Linear Algebra
Axis	Basis vector
X-axis	$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$
Y-axis	$e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$
Z-axis	$e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$
3-D Euclidean Space	$Span\{e_1, e_2, e_3\}$
Perpendicular	Orthogonal
Dimension	Dimension

Linearly Dependent Vectors

The vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ are called linearly dependent if and only if we can find scalars $c_1, \dots, c_k$ such that $c_1\boldsymbol{v}_1 + \cdots + c_k\,\boldsymbol{v}_k = \mathbf{0}$ and not all c_i, $i = \{1, .., k\}$, are zeros.

ML Use Case:

Dependency (collinearity) implies redundancy in data

Because:

$$\boldsymbol{v}_1 = -\frac{c_2}{c_1}\boldsymbol{v}_2 - \cdots \frac{c_k}{c_1}\boldsymbol{v}_k \quad \text{(assume } c_1 \neq 0\text{)}$$

Example 1: Linearly Dependent Vectors

- Given $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\boldsymbol{v}_2 = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$.

Are $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ linearly dependant?

Solution: Yes

$$-5 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} 5 \\ 10 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{or } \boldsymbol{v}_2 = 5\, \boldsymbol{v}_1$$

$$c_1 = -5\,, c_2 = 1, and\, c_1 \boldsymbol{v}_1 + c_2\, \boldsymbol{v}_2 = \boldsymbol{0}$$

Example 2: Linearly Dependent Vectors

- Given $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 3 \\ 6 \\ -3 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix}, \boldsymbol{v}_4 = \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix}$.

Are $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3, and\ \boldsymbol{v}_4$ linearly dependent?

Solution: Yes.

$$c_1 = -3, \ c_2 = 1, c_3 = 0, and \ \ c_4 = 0$$

$$-3 * \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + 1 * \begin{bmatrix} 3 \\ 6 \\ -3 \end{bmatrix} + 0 * \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix} + 0 * \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Linearly Independent Vectors

The vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are called linearly independent if the following condition is satisfied:

$c_1\boldsymbol{v}_1 + \cdots + c_k\,\boldsymbol{v}_k = \boldsymbol{0}$ if and only if $c_1 = c_2 = .. = c_k = 0$

- Maximum number of independent vectors indicate dimension of the space

Example: Linearly Independent Vectors

- Given $\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix}$. Are these linearly independent?

Answer: Yes, $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3$=0 only if $c_1 = c_2 = c_3 = 0$

- Given $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Are these linearly independent?

Answer: Yes, $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3$=0 only if $c_1 = c_2 = c_3 = 0$

Home Work: Checking dependency

- Read **Section 1.2 Row Reduction** of
  https://textbooks.math.gatech.edu/ila/ila.pdf

- Key Idea:
  - Treat the vectors as columns of a matrix.
  - Convert the matrix to row reduced echelon form.
  - If the row reduced echelon form contains an all zero row, then the vectors are linearly dependent.

Span

- Let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ be vectors. Then span of $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ is the collection of all linear combinations of these vectors, i.e.,

$$Span\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k\}$$
$$= \{x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_k\boldsymbol{v}_k | x_1, x_2, \dots, x_k \in \mathbb{R}\}$$

Set builder notation:
you will learn notations of sets tomorrow

Span: n-D Space

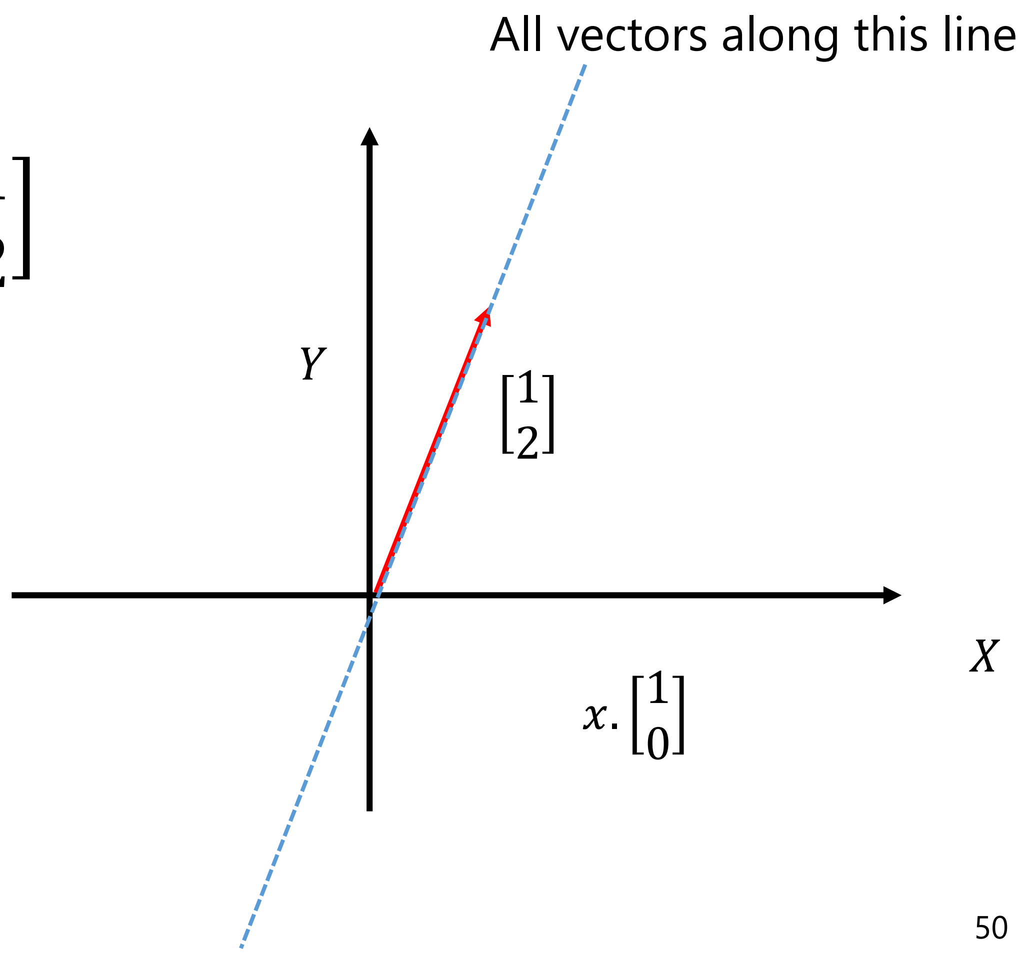
- Every n-D vector can be considered as a linear combination of $\boldsymbol{e}_i, i = \{1, \dots, n\}$, where

$$\boldsymbol{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} 0 \\ \\ i-1 \\ i \\ i+1 \\ \\ n-1 \end{array}$$

Span: Examples

- What is the $\text{span}\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$?

$$\text{span}\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\} = c_1\begin{bmatrix}1\\2\end{bmatrix}$$

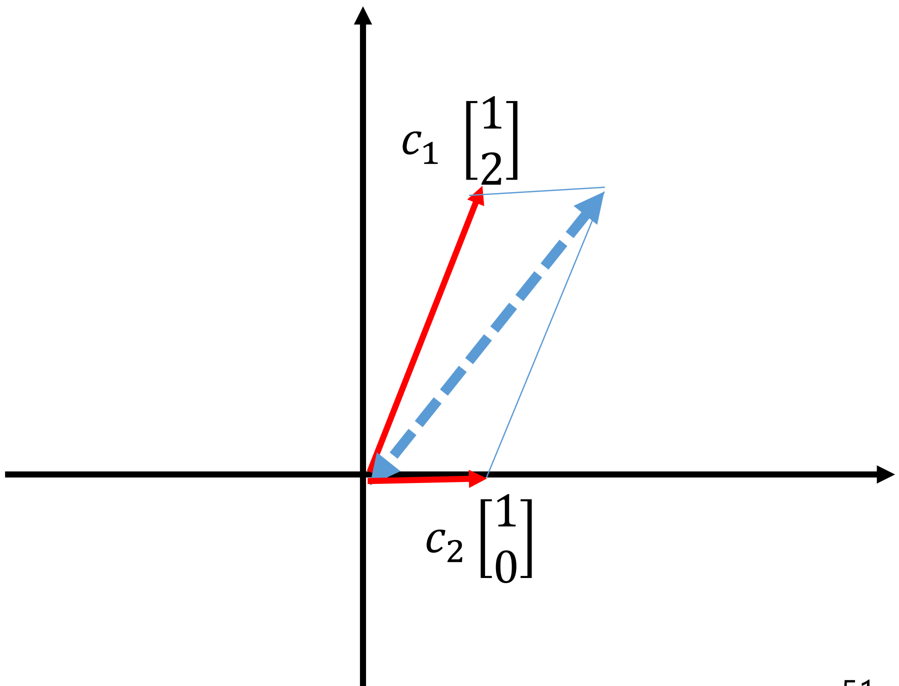


Span: Examples

- What is the span $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}\right\}$?

$$\text{span}\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\} = c_1\begin{bmatrix}1\\2\end{bmatrix} + c_2\begin{bmatrix}1\\0\end{bmatrix}$$

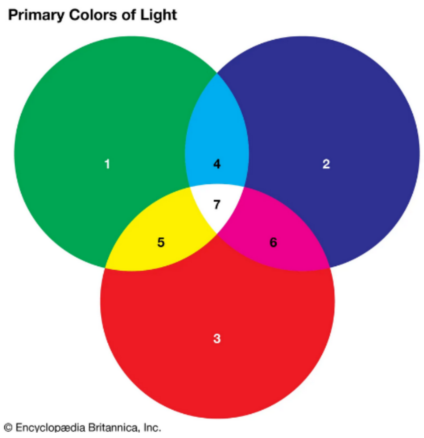
All vectors in 2D space



Basis

- A set of vectors are called basis for a space if they satisfy the following conditions
  - ➢ The vectors are **linearly independent**
  - ➢ The vectors **span** the **entire space**
- A basis is a maximal independent set, i.e., we cannot add vectors to a set without losing independence

Analogy: Color Mixing



- Primary (Basis) Colors: Red, Green, and Blue

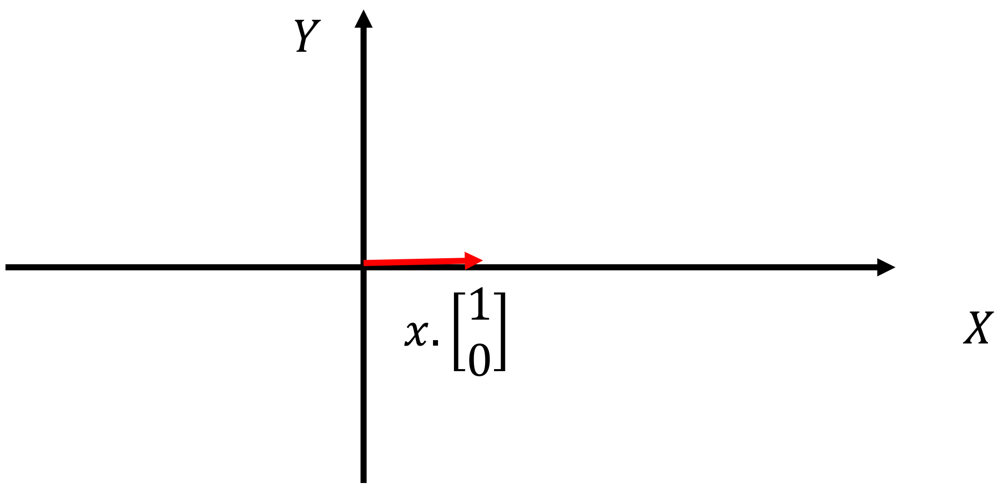
- The three primary colors span the color space
  - $Span\{Red, Green, Blue\} = Color\ Space$

- Red, Green and Blue are linearly independent or non-redundant colors

- Other colors are linear combinations of these primary colors

Basis Vectors: Examples

- Can the set of vectors $\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\}$ be the basis for the 2D space?
  - ➢No, it can only span X-axis



Basis Vectors: Examples

- Can the set of vectors $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$ be the basis for 2D space?
  - Yes
    - The set of vectors $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$ are linearly independent
    - They span the 2D space, i.e., any vector $\begin{bmatrix}x_1\\x_2\end{bmatrix}$ in 2D space can be expressed as linear combination of the set of vectors $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$

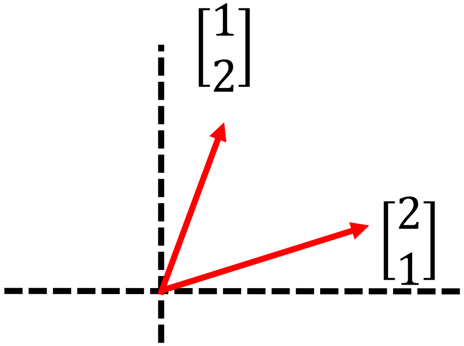
$$\begin{bmatrix}x_1\\x_2\end{bmatrix} = c_1\begin{bmatrix}1\\0\end{bmatrix} + c_2\begin{bmatrix}0\\1\end{bmatrix}$$

$$c_1 = x_1, c_2 = x_2$$

Basis Vectors: Examples

- Can the set of vectors $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}\right\}$ be the basis for 2D space?
    - Yes
        - The set of vectors $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}\right\}$ are linearly independent
        - They span the 2D space, i.e., any vector $\begin{bmatrix}x_1\\x_2\end{bmatrix}$ in 2D space can be expressed as linear combination of the set of vectors $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}\right\}$

$$\begin{bmatrix}x_1\\x_2\end{bmatrix} = c_1\begin{bmatrix}1\\2\end{bmatrix} + c_2\begin{bmatrix}2\\1\end{bmatrix}$$

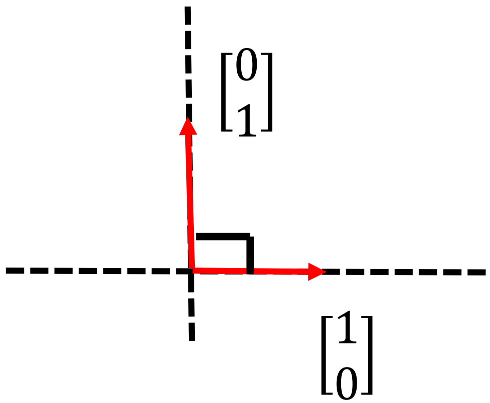


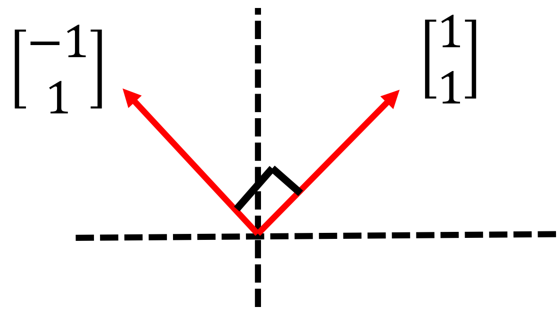
Basis Vectors: Examples

- Can the set of vectors $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}\right\}$ be the basis for 2D space?
  - No.
    - The set of vectors $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}\right\}$ linearly dependent
      - We can express one of the vectors as linear combination of other vectors.
        - How?
    - They span the 2D space

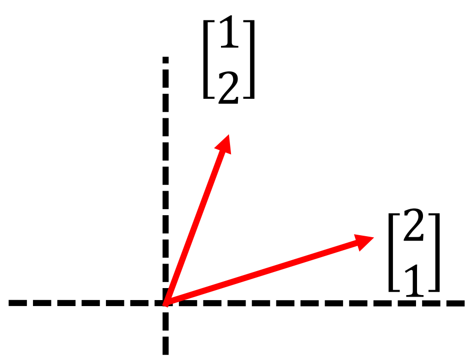
Basis vectors need not be orthogonal

Orthogonal Basis Vectors





Non-Orthogonal Basis Vectors



Dimension

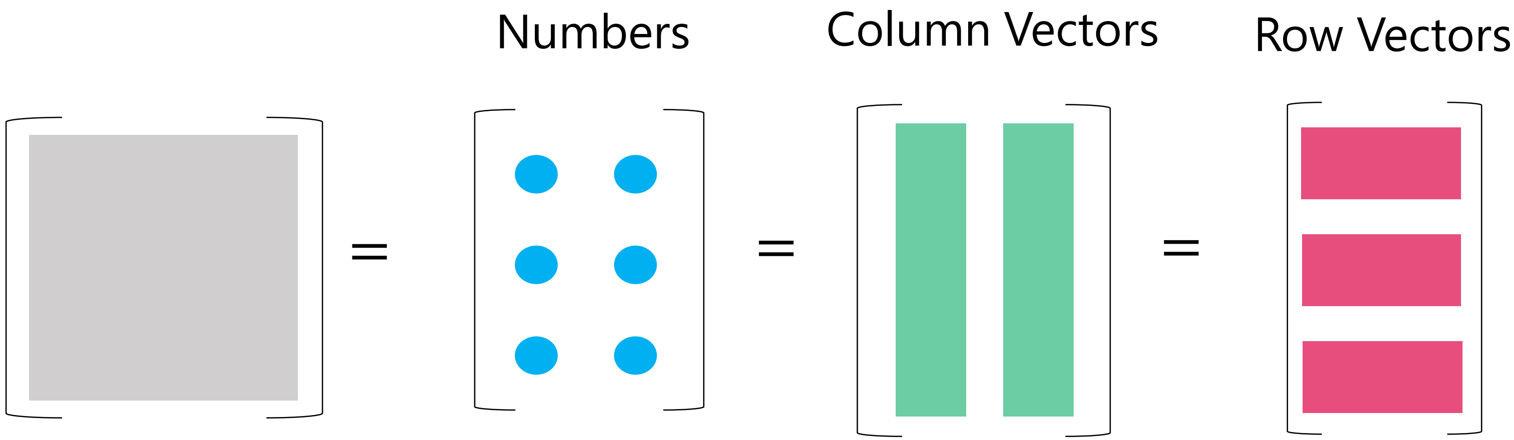
- Number of vectors in bases is called dimension
  - Indicates "degrees of freedom" of space

- Any two bases of a space contain same number of vectors
  - Example: $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$ and $\left\{\begin{bmatrix}1\\2\end{bmatrix},\begin{bmatrix}2\\1\end{bmatrix}\right\}$ for 2D space

Part II: Matrices

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

How to look at matrices?



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

1 matrix

6 numbers

2 Column Vectors with 3 numbers

3 Row Vectors with 2 numbers

https://github.com/kenjihiranabe/The-Art-of-Linear-Algebra

Matrices: Agenda

- Types of Matrices
- Matrix Multiplication
- Transpose and Inverse
- Determinant
- Eigen Value
- Eigen Vectors

Size and Entries of a Matrix

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \left(x_{ij}\right)_{3\times 3} = \left(x_{ij}\right)$$

(3,2)-th element
Element in 3rd row and 2nd column

- Size of the matrix X is $m \times n$
- X is an $m \times n$ matrix
- x_{ij} is the (i, j)-th entry of the matrix X

Square Matrices

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

non-diagonal entries of $\boldsymbol{A}$

diagonal of $\boldsymbol{A}$

- $\boldsymbol{A}$ is an $n \times n$ matrix (or) $\boldsymbol{A}$ is a square matrix of order n
- Square matrix has same number of rows and columns
- $a_{11}, a_{22}, \dots, a_{nn}$ are called diagonal entries
- $a_{ij,}\ i \neq j$ are called off-diagonal entries

Types of Square Matrices

- Diagonal Matrix:
  - Off-diagonal elements are zero
  - $a_{ij} = 0, \; when \; i \neq j$

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

- Scalar Matrix:
  - Diagonal matrix with same diagonal values
  - $a_{ij} = \begin{cases} 0 & for \quad i \neq j \\ c, & for \quad i = j \end{cases}$

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

- Identity Matrix ($\boldsymbol{I}_n$)
  - Diagonal matrix with all diagonal entries are equal to 1
  - $a_{ij} = \begin{cases} 0 & for \quad i \neq j \\ 1, & for \quad i = j \end{cases}$

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

Types of Square Matrices

- **Zero Matrix** ($\mathbf{0}_{m\times n}$):
  - Can be non-square
  - $a_{ij} = 0\ \forall i, j$.
- **Symmetric Matrix**:
  - k-th row is equal to k-th column, for all k
  - $a_{ij} = a_{ji}\ \forall i, j$
- **Upper Triangular Matrix:**
  - All entries below diagonal are zero
  - $a_{ij} = 0\ \ \forall i > j$
- **Lower Triangular Matrix:**
  - All entries above diagonal are zero
  - $a_{ij} = 0\ \forall i < j$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

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

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

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

Matrix Arithmetic

- **Matrix Equality** $(\boldsymbol{A} = \boldsymbol{B})$:
  - Both matrices must have same size
  - $a_{ij} = b_{ij}$ for all i, j

- **Matrix Addition** $(\boldsymbol{A} + \boldsymbol{B})$:
  - Both matrices must have same size
  - Add the corresponding entries, i.e., $(a_{ij} + b_{ij})_{m\times n}$

- **Matrix Subtraction** $(\boldsymbol{A} - \boldsymbol{B})$:
  - Both matrices must have same size
  - Subtract the corresponding entries, i.e., $(a_{ij} - b_{ij})_{m\times n}$

- **Scalar Multiplication** $(\mathrm{c}\boldsymbol{A})$:
  - Multiply every entry of $\boldsymbol{A}$ by c, i.e., $(ca_{ij})_{m\times n}$

- **Negative of a matrix** $(-\boldsymbol{A})$:
  - Negative of each element of $\boldsymbol{A}$, i.e., $(-a_{ij})_{m\times n}$

Properties of Matrix Addition

- Matrix addition is commutative

$$\boldsymbol{A} + \boldsymbol{B} \; = \; \boldsymbol{B} + \boldsymbol{A}$$

- Matrix addition is associative

$$\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \; \boldsymbol{B}) + \boldsymbol{C}$$

Matrix Multiplication

- We can multiply two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ as $\boldsymbol{AB}$ only if the number of columns of $\boldsymbol{A}$ are equal to the number of rows of $\boldsymbol{B}$

Size of A	**Size of B**	**Size of AB**
$m \times n$	$n \times p$	$m \times p$

Must be equal

Matrix Multiplication

$\boldsymbol{A}$	$\boldsymbol{B}$	$\boldsymbol{AB}$	$\boldsymbol{BA}$
2×3	3×4	2×4	Not Possible
2×3	3×2	2×2	3×3
4×1	2×4	Not Possible	2×1
1×3	2×4	Not Possible	Not Possible

Properties of Matrix Multiplication

- Matrix multiplication is not commutative

$$\boldsymbol{AB} \neq \boldsymbol{BA}$$

- Matrix multiplication is associative

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$$

- Matrix multiplication obeys distributive law

$$\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$$

Properties of Matrix Multiplication

- Zero matrix behaves like number ‘0’
  - $\boldsymbol{A}_{m\times n}\ \mathbf{0}_{n\times p} = \mathbf{0}_{m\times p}$
  - $\mathbf{0}_{p\times m}\ \boldsymbol{A}_{m\times n} = \mathbf{0}_{p\times n}$

- Identity matrix behaves like number ‘1’
  - $\boldsymbol{A}_{m\times n}\ \boldsymbol{I}_{n} = \boldsymbol{A}_{m\times n}$
  - $\boldsymbol{I}_{m}\boldsymbol{A}_{m\times n} = \boldsymbol{A}_{m\times n}$

Powers of a Matrix

- Let $\boldsymbol{A}$ be a square matrix and n be a non-negative integer

- $\boldsymbol{A}^n$ is defined as:
  - $\boldsymbol{A}^n = \underbrace{\boldsymbol{A}\ \boldsymbol{A} \ldots\ \boldsymbol{A}}_{n \text{ times}},\ n \geq 1$
  - $\boldsymbol{A}^0 = \boldsymbol{I}$

Matrix Transpose

- Transpose of a matrix $\boldsymbol{A}$ is denoted as $\boldsymbol{A}^T$
- $\boldsymbol{A}^T$ can be obtained by interchanging the rows and columns of a matrix
- i.e., if $\boldsymbol{A} = (a_{ij})_{m\times n}$, then $\boldsymbol{A}^T = (a_{ji})_{n\times m}$
- **Example**:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$$\boldsymbol{A}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Properties of Transpose

- Let **A** and $\boldsymbol{B}$ be $m \times n$ matrices, and c be a scalar, then
  - $(\boldsymbol{A}^T)^T = \boldsymbol{A}$
  - $(\boldsymbol{A} + \boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$
  - $(\boldsymbol{cA})^T = c\boldsymbol{A}^T$
- Let **A** be an $m \times n$ matrix and B be an $n \times p$ matrix, then
  - $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$

Symmetric Matrix

- A square matrix is symmetric if $\boldsymbol{A}^T = \boldsymbol{A}$

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

- The matrices $\boldsymbol{A}^T\boldsymbol{A}$ and $\boldsymbol{A}\boldsymbol{A}^T$ are symmetric

Inverse of a Matrix

- Let $\boldsymbol{A}$ be a square matrix of order n. Matrix $\boldsymbol{A}$ is invertible if there exists a matrix $\boldsymbol{B}$ such that $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$

- Matrix $\boldsymbol{B}$ is called inverse of matrix $\boldsymbol{A}$,

i.e., $\boldsymbol{B} = \boldsymbol{A}^{-1}$

Inverses and Matrix Operations

- Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be invertible matrices and c be a scalar

  - $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$

  - $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$

  - $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$

  - $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

Determinant

- A function that maps a square matrix to a scalar

- Let $\boldsymbol{A}$ be a square $n \times n$ matrix. Let $\boldsymbol{A}_{ij}$ be a matrix obtained by removing the i-th row and j-th column from $\boldsymbol{A}$. The determinant of $\boldsymbol{A}$ or $|\boldsymbol{A}|$ is

$$|\boldsymbol{A}| = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \left|\boldsymbol{A}_{1j}\right|$$

Determinant: Example

- **Example**: $\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

$$|\boldsymbol{A}| = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \left|\boldsymbol{A}_{1j}\right|$$

$$|A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} \square & \square & \square \\ \square & e & f \\ \square & h & i \end{vmatrix} - b \begin{vmatrix} \square & \square & \square \\ d & \square & f \\ g & \square & i \end{vmatrix} + c \begin{vmatrix} \square & \square & \square \\ d & e & \square \\ g & h & \square \end{vmatrix}$$

$$= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= aei + bfg + cdh - ceg - bdi - afh.$$

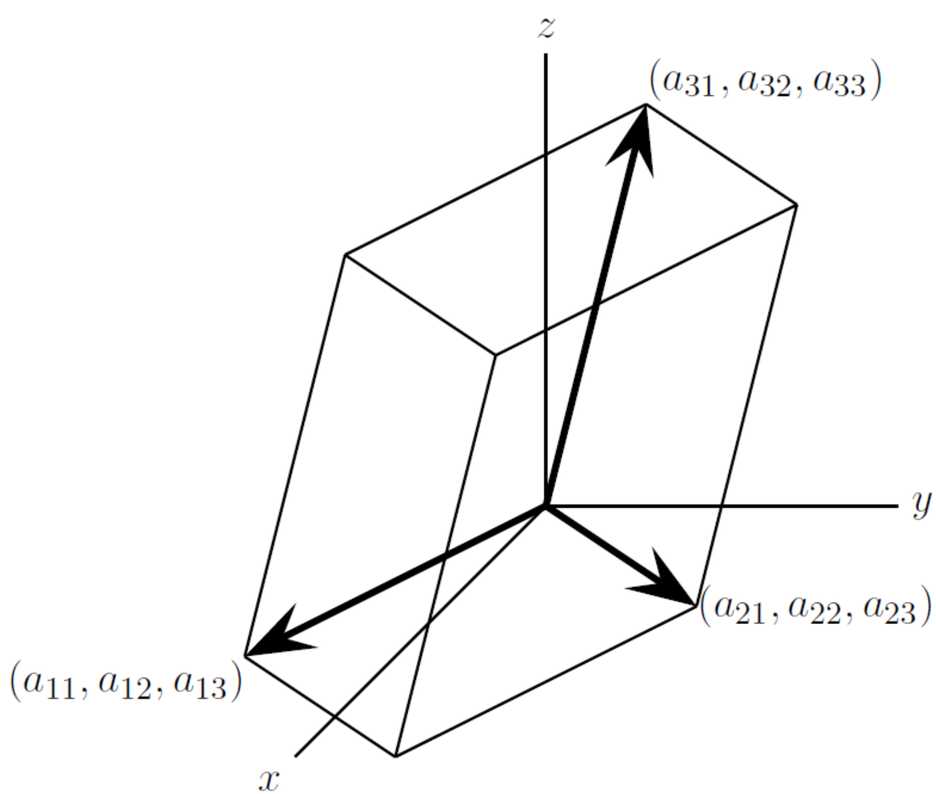
Determinant: Example

- Find the determinant of $\boldsymbol{A} = \begin{bmatrix} 2 & 1 & 3 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{bmatrix}$

- Answer: 15

Determinants and Volumes

- Absolute value of determinant of A indicates the volume of the parallelepiped formed by the rows (or columns) of matrix $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$



Homework: Draw the parallelepiped for identity matrix

Properties of Determinant

- If a matrix is diagonal/triangular, its determinant is product of diagonal elements

- Example:

$$\boldsymbol{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}, |\boldsymbol{A}| = 2 * 3 * 4 = 24$$

$$\boldsymbol{B} = \begin{bmatrix} 2 & 1 & 10 \\ 0 & 3 & 20 \\ 0 & 0 & 4 \end{bmatrix}, |\boldsymbol{B}| = 2 * 3 * 4 = 24$$

Properties of Determinant

- If two rows/columns of a matrix are equal, then its determinant is zero

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}, |\boldsymbol{A}| = 0$$

- Or, in general, if a matrix contains linearly dependent rows/columns, its determinant is zero.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 2a_{11} & 2a_{12} & 2a_{13} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}, |\boldsymbol{A}| = 0$$

Homework: How does the parallelepiped looks like for this case?

Properties of Determinant

- For any two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, the determinant of the product $\boldsymbol{AB}$ is the product of the determinants of matrices A and B,

$$i.e., |\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}|$$

- Determinant of a transpose matrix is the same as determinant of the original matrix,

$$\text{i.e., } |\boldsymbol{A}^T| = |\boldsymbol{A}|$$

Singular Matrix

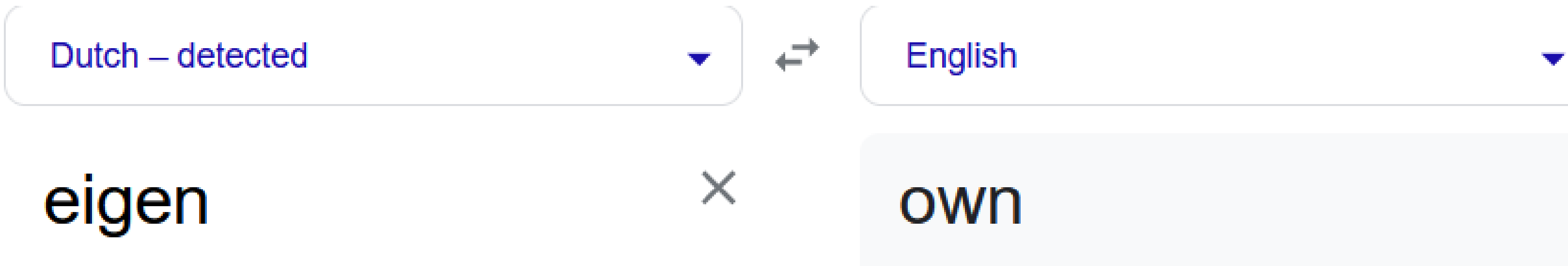
- If the determinant of a matrix is zero, then it's a singular matrix

- Example:

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 & 10 \\ 1 & 3 & 20 \\ 4 & 2 & 20 \end{bmatrix} \text{ is a singular matrix.}$$

- A matrix is invertible if, and only if its determinant is non-zero

Eigenvalues and Eigenvectors



Let $\boldsymbol{A}$ be a $n \times n$ square matrix. Let λ be an eigenvalue (scalar) and $\boldsymbol{v} \neq \mathbf{0}$ be an eigenvector of the matrix $\boldsymbol{A}$, then $\boldsymbol{Av} = \lambda \boldsymbol{v}$

How to find Eigenvalues and Eigenvectors?

- Compute the determinant of $\boldsymbol{A} - \lambda \boldsymbol{I}$
  - The determinant is a polynomial of degree n, where n is the # of rows/columns of matrix $\boldsymbol{A}$

- Find the roots of the polynomials $|\boldsymbol{A} - \lambda \boldsymbol{I}|$.
  - The roots of the degree n polynomial are the eigenvalues of the matrix $\boldsymbol{A}$

- For each eigen value solve the equation $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}=0$. The solutions of this equation are the eigenvectors of $\boldsymbol{A}$

Chapter 5 of https://textbooks.math.gatech.edu/ila/ila.pdf

Eigenvalues and Eigenvectors

Eigenvalues of a diagonal and triangular matrices are its diagonal elements

Examples:

Eigenvalues of $\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$ are $a_{11}, a_{22}, and\ a_{33}$

Eigenvalues of $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$ are $a_{11}, a_{22}, and\ a_{33}$

Determinant and Eigenvalues

- Determinant of a matrix is equal to product of eigenvalues

- Example:

$$\boldsymbol{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}, |\boldsymbol{A}| = 2 * 3 * 4 = 24$$

$$\boldsymbol{B} = \begin{bmatrix} 2 & 1 & 10 \\ 0 & 3 & 20 \\ 0 & 0 & 4 \end{bmatrix}, |\boldsymbol{B}| = 2 * 3 * 4 = 24$$

Eigenvectors of Symmetric Matrices

- Eigenvectors of symmetric matrices are orthogonal to each other

- **Example**:

  $A^T A$ and AA^T are symmetric

  Eigenvectors of these matrices are orthogonal

Symmetric matrices have orthogonal eigenvectors

Proof: Let $\boldsymbol{A}^{\boldsymbol{T}} = \boldsymbol{A}$ be symmetric matrix with distinct eigenvalues

Let $\boldsymbol{Ax} = \lambda\boldsymbol{x}$ and $\boldsymbol{Ay} = \alpha\boldsymbol{y}$ and $\lambda \neq \alpha$

1. Take transpose on both sides of $\boldsymbol{Ax} = \lambda\boldsymbol{x}$

$$(\boldsymbol{Ax})^{\boldsymbol{T}} = (\lambda\boldsymbol{x})^{\boldsymbol{T}}$$

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}^{\boldsymbol{T}} = \lambda\boldsymbol{x}^{\boldsymbol{T}}$$

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A} = \lambda\boldsymbol{x}^{\boldsymbol{T}}$$

2. Multiply both sides of $\boldsymbol{Ay} = \boldsymbol{\alpha y}$ by $\boldsymbol{x}^{\boldsymbol{T}}$

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{Ay} = \boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{\alpha y}$$

$$\lambda\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{y} = \boldsymbol{\alpha}\,\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{y}$$

Since $\lambda \neq \alpha$, $\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{y}$ must be zero

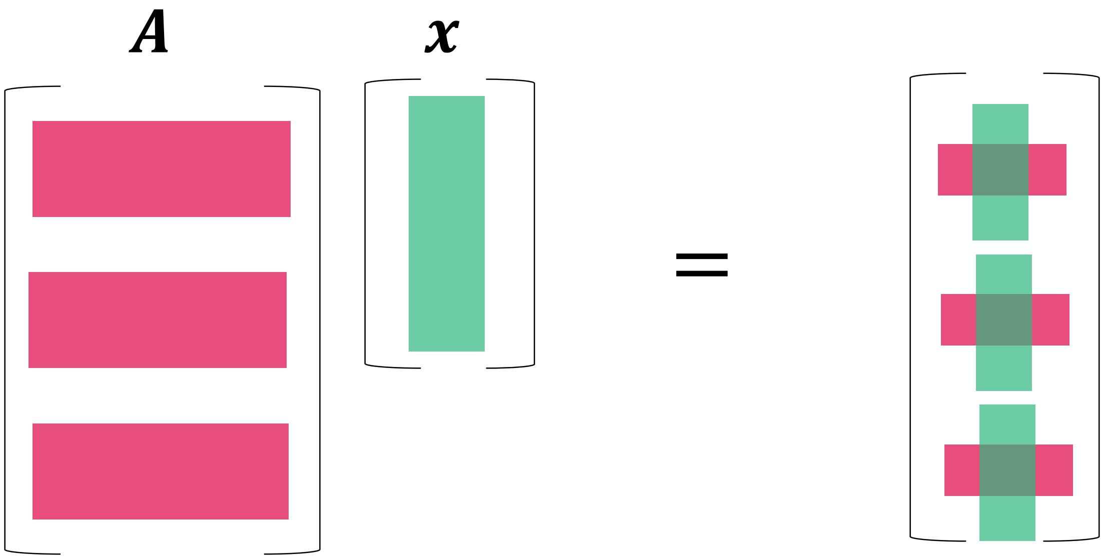
$\therefore$ $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal

Product of Matrix and Vector

- $\boldsymbol{Ax}$: Product of Matrix ($\boldsymbol{A}$) and Column Vector ($\boldsymbol{x}$)
  - $\boldsymbol{Ax}1$: Dot-product view
  - $\boldsymbol{Ax}2$: Linear combination view
- $\boldsymbol{yA}$: Product of Row Vector ($\boldsymbol{y}$) and Matrix ($\boldsymbol{A}$)
  - $\boldsymbol{yA}1$: Dot-product view
  - $\boldsymbol{yA}2$: Linear combination view

https://github.com/kenjihiranabe/The-Art-of-Linear-Algebra

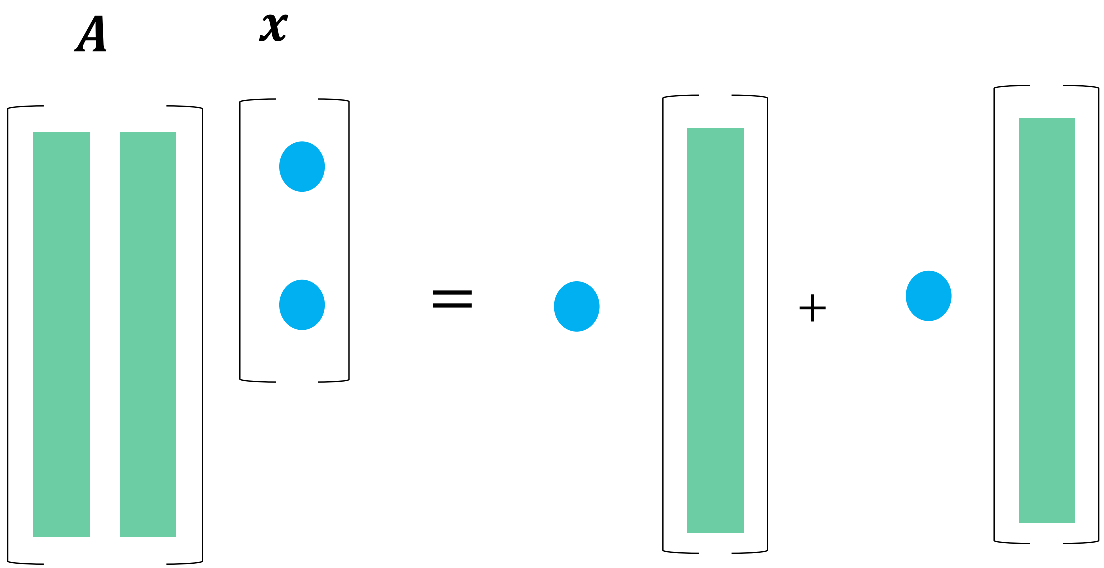
$\boldsymbol{Ax}$1: Product of Matrix $\boldsymbol{A}$ and Column Vector $\boldsymbol{x}$



The row vectors of $\boldsymbol{A}$ are multiplied by a column vector $\boldsymbol{x}$ and become the three dot-product elements of $\boldsymbol{Ax}$.

$$\boldsymbol{Ax} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (x_1 + 2x_2) \\ (3x_1 + 4x_2) \\ (5x_1 + 6x_2) \end{bmatrix}$$

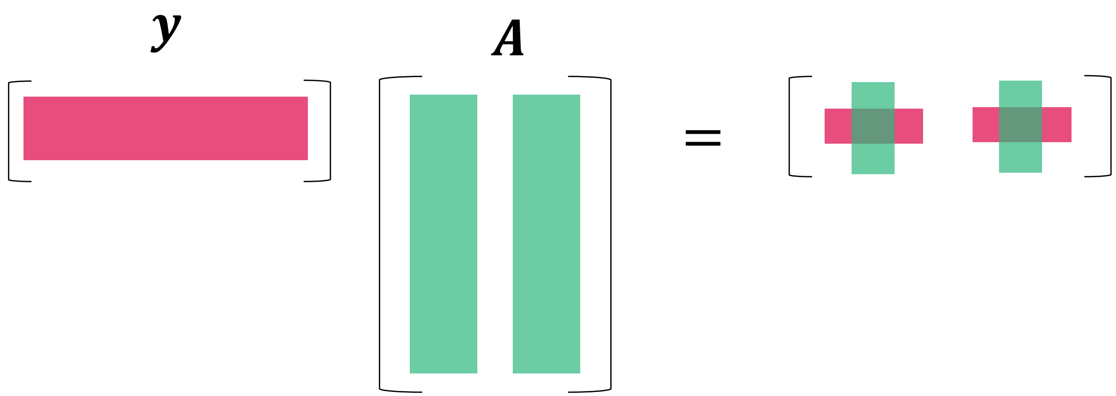
$\boldsymbol{Ax}$2: Product of Matrix $\boldsymbol{A}$ and Column Vector $\boldsymbol{x}$



The product $\boldsymbol{Ax}$ is a linear combination of the column vectors of $\boldsymbol{A}$.

$$\boldsymbol{Ax} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$$

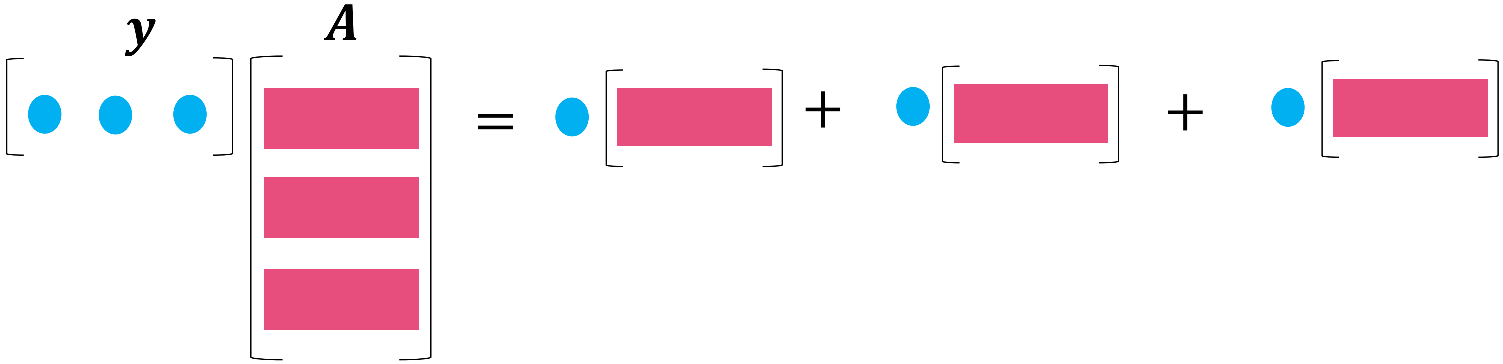
$\boldsymbol{yA}1$: Product of a Row Vector $\boldsymbol{y}$ and Matrix $\boldsymbol{A}$



$$\boldsymbol{yA} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} (y_1 + 3y_2 + 5y_3) & (2y_1 + 4y_2 + 6y_3) \end{bmatrix}$$

A row vector $\boldsymbol{y}$ is multiplied by the two column vectors of $\boldsymbol{A}$ and become the two dot-product elements of $\boldsymbol{yA}$.

$\boldsymbol{yA}$2: Product of a Row Vector $\boldsymbol{y}$ and Matrix $\boldsymbol{A}$



$$\boldsymbol{yA} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = y_1\begin{bmatrix} 1 & 2 \end{bmatrix} + y_2\begin{bmatrix} 3 & 4 \end{bmatrix} + y_3\begin{bmatrix} 5 & 6 \end{bmatrix}$$

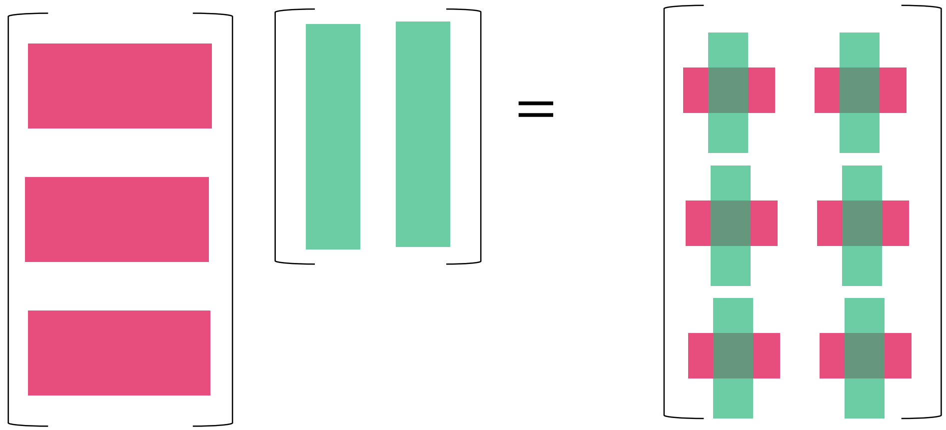
The product $\boldsymbol{yA}$ is a linear combination of the row vectors of $\boldsymbol{A}$.

Product of Matrix and Matrix

- $\boldsymbol{MM}$: Product of two matrices
  - ➢ $\boldsymbol{MM}1$: Dot-product view
  - ➢ $\boldsymbol{MM}2$: Outer Product view
  - ➢ $\boldsymbol{MM}3$: Linear combination of Rows view
  - ➢ $\boldsymbol{MM}4$: Linear combination of Columns view

https://github.com/kenjihiranabe/The-Art-of-Linear-Algebra

$\boldsymbol{MM}1$: Dot Product Interpretation

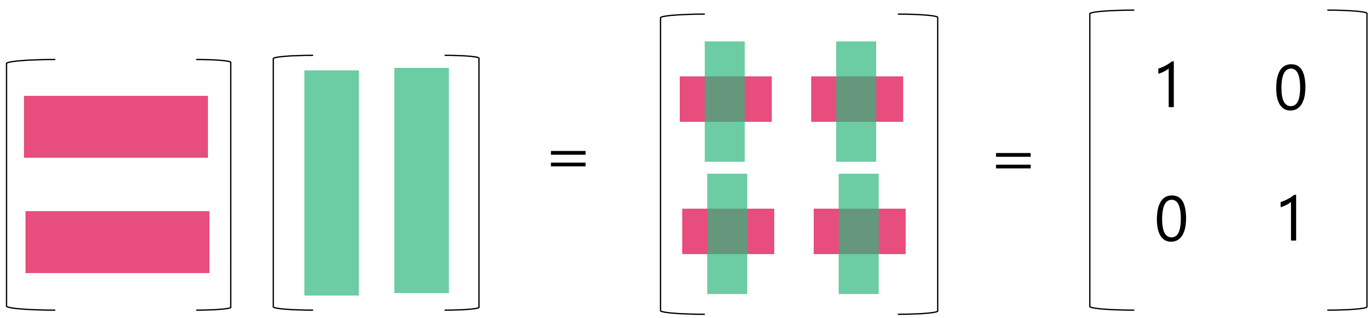


Every element becomes a dot product of row vector and column vector.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} (b_{11}+2b_{21}) & (b_{12}+2b_{22}) \\ (3b_{11}+4b_{21}) & (3b_{12}+4b_{22}) \\ (5b_{11}+6b_{21}) & (5b_{12}+6b_{22}) \end{bmatrix}$$

Orthogonal Matrix

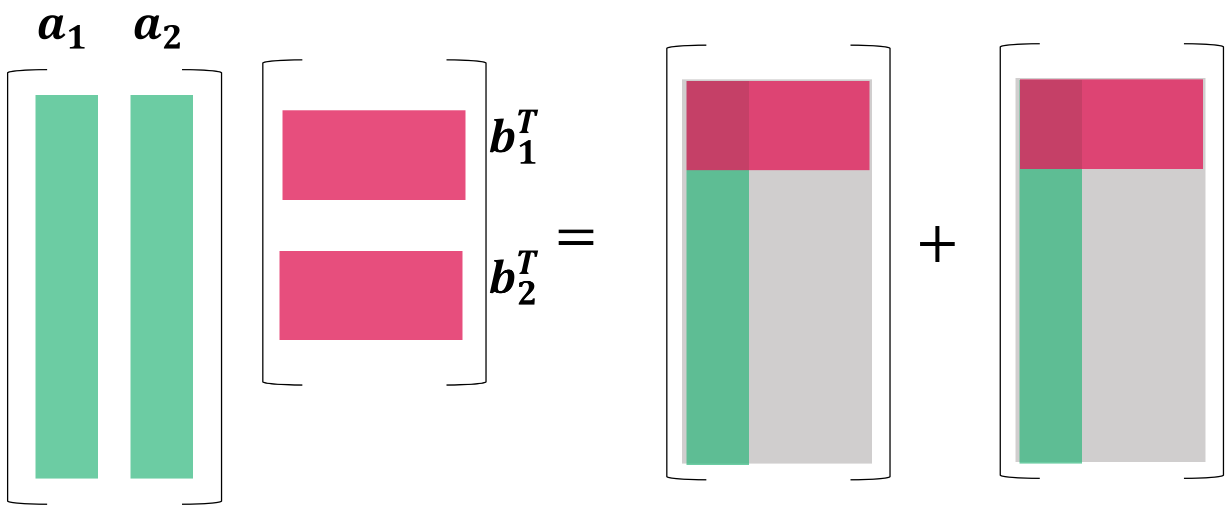
- A matrix $\boldsymbol{Q}$ is called orthogonal matrix if $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$, where $\boldsymbol{I}$ is an identity matrix



- How to construct a orthogonal matrix?
  - ➢ Select a set of orthonormal vectors as columns, i.e., $\boldsymbol{Q} = [\boldsymbol{q_1} \quad \boldsymbol{q_2} \cdots \boldsymbol{q_n}]$

$$\begin{bmatrix} \boldsymbol{q_1}^T \\ \boldsymbol{q_2}^T \end{bmatrix} [\boldsymbol{q_1} \quad \boldsymbol{q_2}] = \begin{bmatrix} \boldsymbol{q_1}^T\boldsymbol{q_1} & \boldsymbol{q_1}^T\boldsymbol{q_2} \\ \boldsymbol{q_2}^T\boldsymbol{q_1} & \boldsymbol{q_2}^T\boldsymbol{q_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$\boldsymbol{MM}2$: Outer Product Interpretation



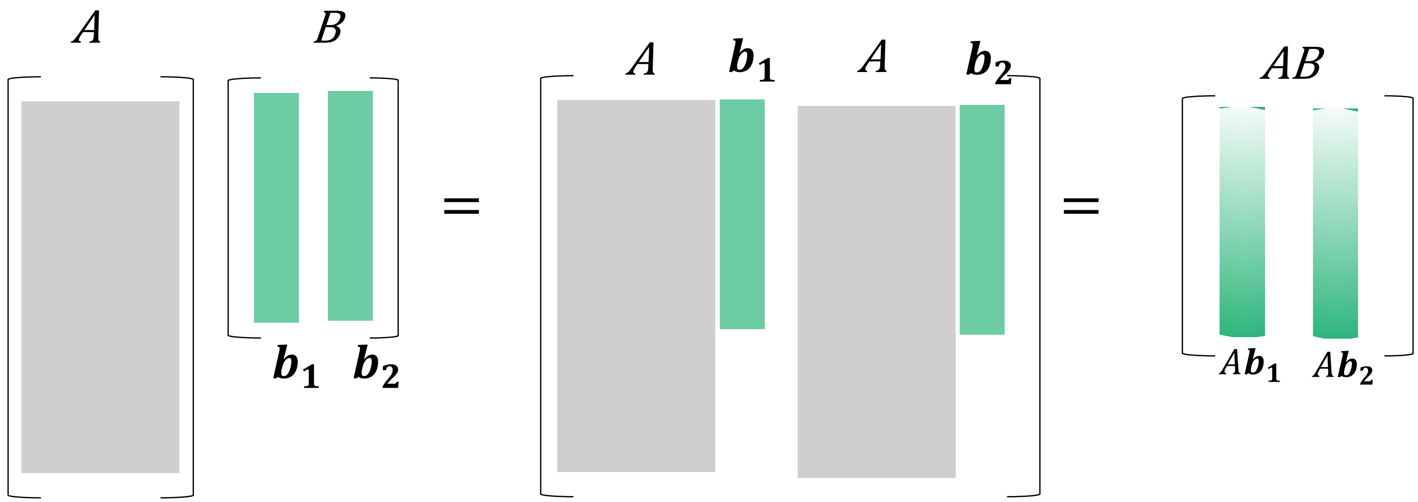
Multiplication AB is broken down to a sum of rank 1 matrices.

$$[\boldsymbol{a}_1 \quad \boldsymbol{a}_2]\begin{bmatrix}\boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T\end{bmatrix} = \boldsymbol{a}_1\boldsymbol{b}_1^T + \boldsymbol{a}_2\boldsymbol{b}_2^T$$

$$\begin{bmatrix}1 & 2 \\ 3 & 4 \\ 5 & 6\end{bmatrix}\begin{bmatrix}b_{11} & b_{12} \\ b_{21} & b_{22}\end{bmatrix} = \begin{bmatrix}1 \\ 3 \\ 5\end{bmatrix}[b_{11} \quad b_{12}] + \begin{bmatrix}2 \\ 4 \\ 6\end{bmatrix}[b_{21} \quad b_{22}]$$

$$= \begin{bmatrix}b_{11} & b_{12} \\ 3b_{11} & 3b_{12} \\ 5b_{11} & 5b_{12}\end{bmatrix} + \begin{bmatrix}2b_{21} & 2b_{22} \\ 4b_{21} & 4b_{22} \\ 6b_{21} & 6b_{22}\end{bmatrix}$$

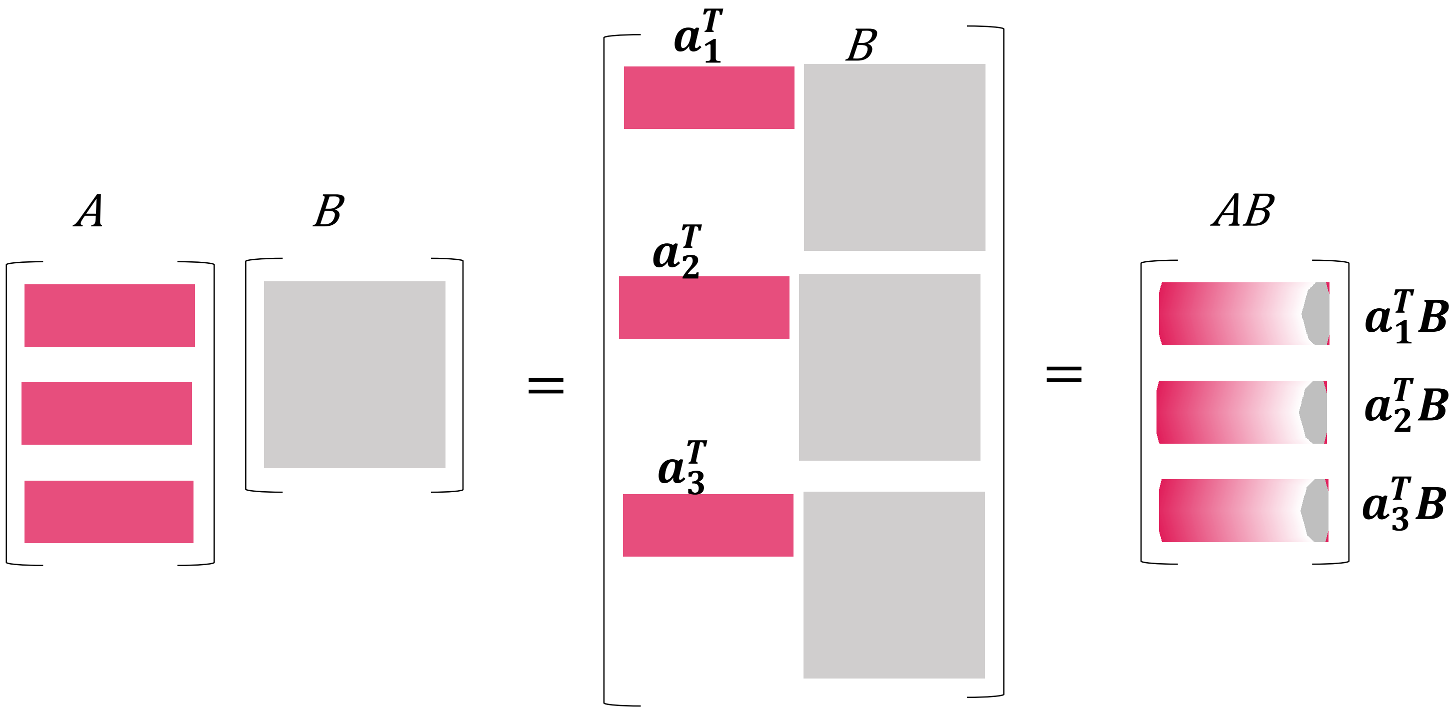
$\boldsymbol{MM}3$: As a Linear Combination of Columns



$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = A[\boldsymbol{b_1} \quad \boldsymbol{b_2}] = [A\boldsymbol{b_1} \quad A\boldsymbol{b_2}]$$

$A\boldsymbol{b_1}$ and $A\boldsymbol{b_2}$ are linear combinations of columns of A.

$\boldsymbol{MM}4$: As a Linear Combination of Rows



$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ \boldsymbol{a}_3^T \end{bmatrix} B = \begin{bmatrix} \boldsymbol{a}_1^T B \\ \boldsymbol{a}_2^T B \\ \boldsymbol{a}_3^T B \end{bmatrix}$$

The rows $\boldsymbol{a}_1^T B$, $\boldsymbol{a}_2^T B$, and $\boldsymbol{a}_3^T B$ are linear combinations of rows of matrix B.

Rank of a Matrix (r)

- Rank of a matrix r indicates number of independent rows/columns of a matrix

- If $\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{m} \times \boldsymbol{n}}$, then $0 \leq r \leq \min(m, n)$

- Only a Zero matrix has a rank of 0

- Number of independent rows of a matrix are equal to number of independent columns

How to find the rank of a matrix?

- Reduce the matrix to row echelon form (ref)

- Number of non-zero rows in the row echelon form indicates the rank of the matrix

Read **Section 1.2 Row Reduction** of
https://textbooks.math.gatech.edu/ila/ila.pdf

Rank-1 Matrix

- Outer product of any two column vectors results in a rank-1 matrix

$$\boldsymbol{x_1} \times \boldsymbol{x_2} = \boldsymbol{x_1}\boldsymbol{x_2^T} = \begin{bmatrix} x_{11} \\ x_{21} \\ \cdot \\ \cdot \\ \cdot \\ x_{n1} \end{bmatrix} \begin{bmatrix} x_{12} & x_{22} & \cdots & x_{n2} \end{bmatrix} = \begin{bmatrix} x_{11}x_{12} & x_{11}x_{22} & \cdots & x_{11}x_{n2} \\ x_{21}x_{12} & x_{21}x_{22} & \cdots & x_{21}x_{n2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ x_{n1}x_{12} & x_{n1}x_{22} & & x_{n1}x_{n2} \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{x_1}x_{12} & \boldsymbol{x_1}x_{22} & \cdots & \boldsymbol{x_1}x_{n2} \end{bmatrix}$$ $\boldsymbol{MM}3$: Only one independent column

$$= \begin{bmatrix} x_{11}\boldsymbol{x_2^T} \\ x_{21}\boldsymbol{x_2^T} \\ \cdot \\ \cdot \\ \cdot \\ x_{n1}\boldsymbol{x_2^T} \end{bmatrix}$$ $\boldsymbol{MM}4$: Only one independent row

Four Spaces of a Matrix

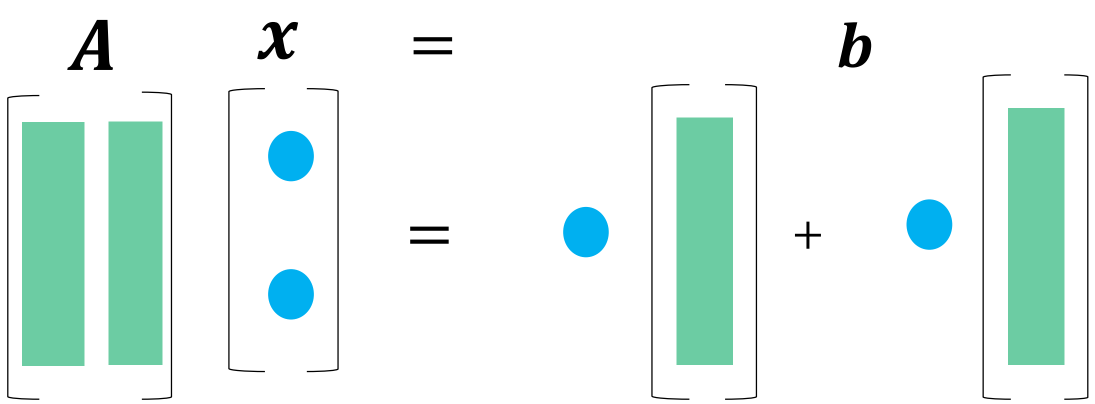
Row Space
Null Space
Column Space
Left Null Space

Column Space of a Matrix

- Space spanned by columns of a matrix $\boldsymbol{A}$

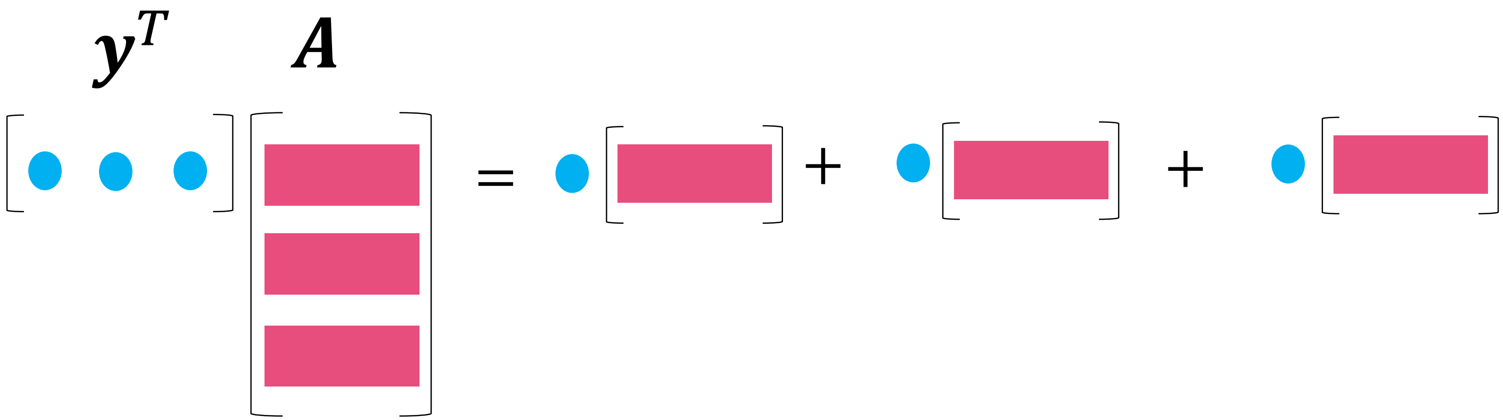
- The set of all vectors $\boldsymbol{b}$ such that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

- Also called range of $\boldsymbol{A}$



Row Space of a Matrix

- Space spanned by rows of a matrix
  - Set of all $\boldsymbol{b}^T$ such that $\boldsymbol{y}^T\boldsymbol{A} = \boldsymbol{b}^T$
  - Called range of $\boldsymbol{A}^T$

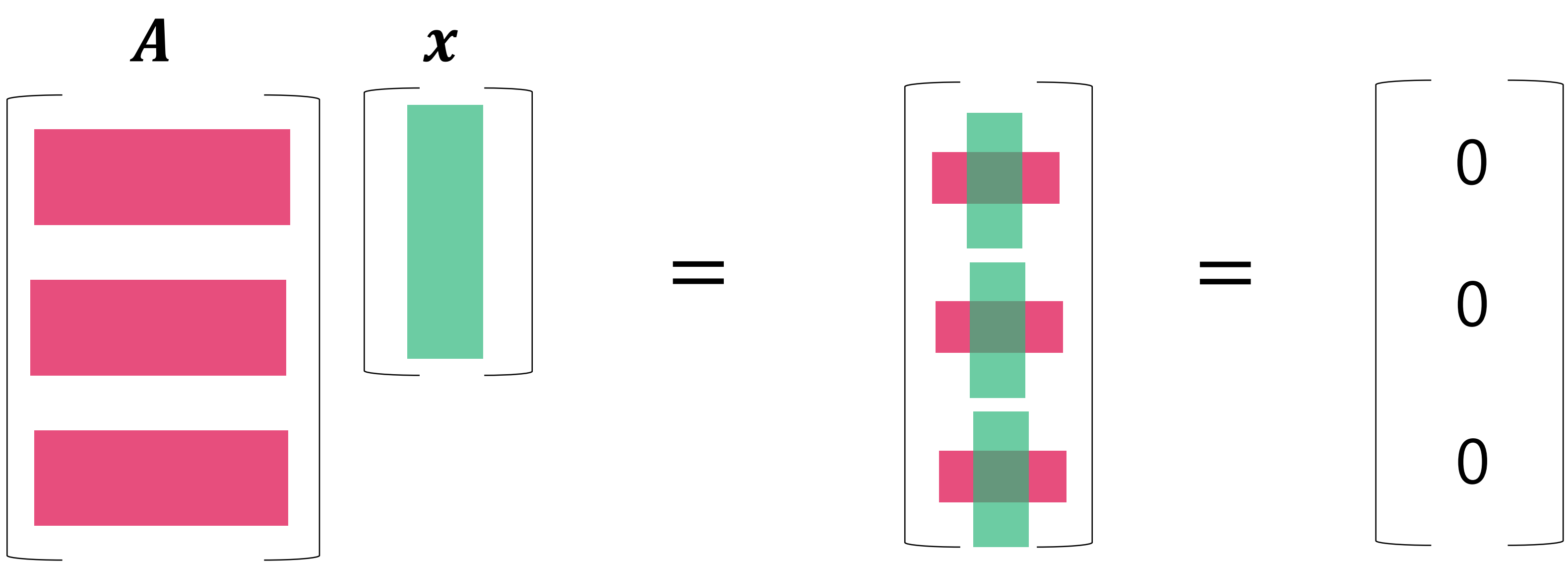


Vectors in row space can be generated by multiplying with row vectors on the left

Null Space of a Matrix

- Set of vectors that span the space that is orthogonal to row space

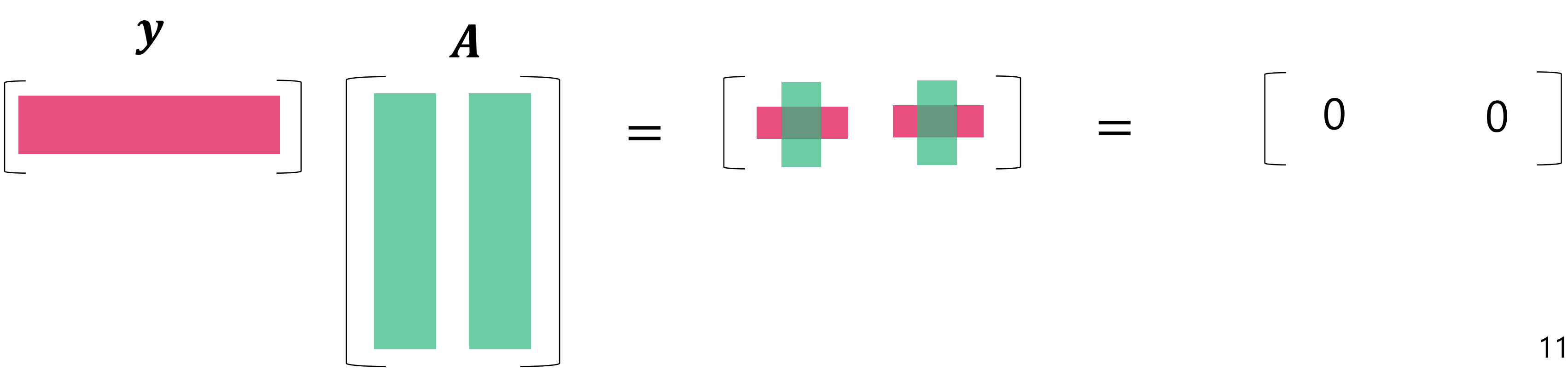
- Set of all vectors $\boldsymbol{x}$ such that $\boldsymbol{Ax} = \boldsymbol{0}$



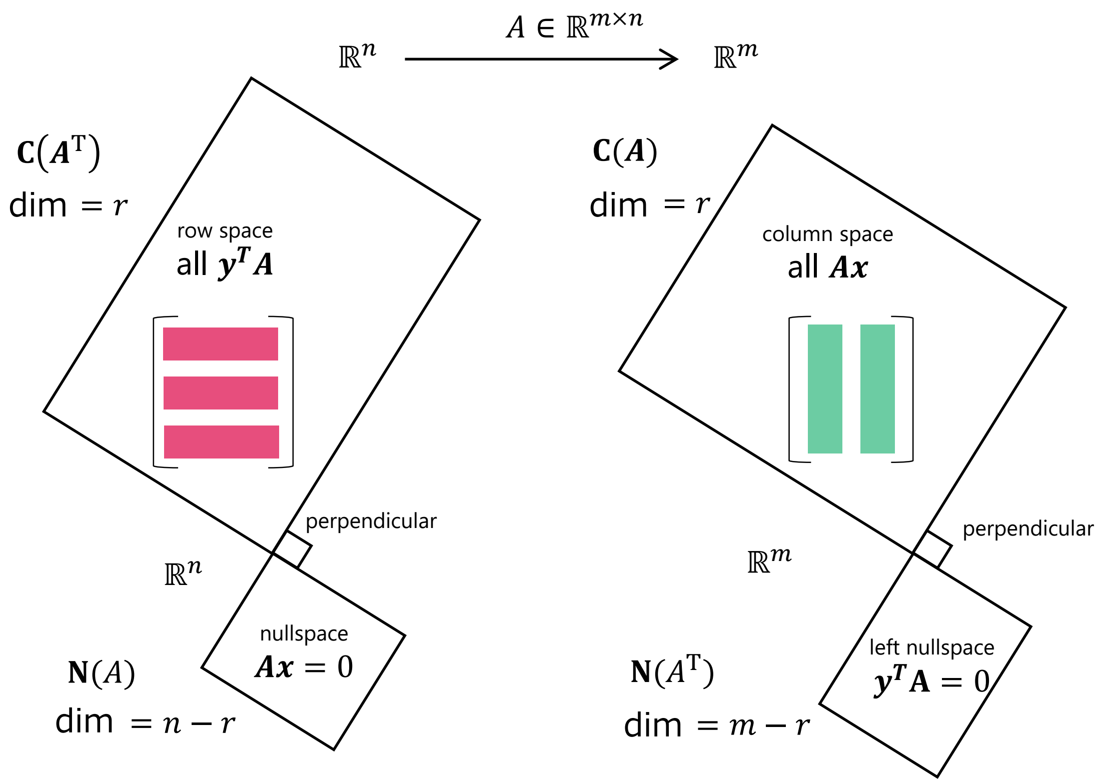
$\boldsymbol{x}$ is orthogonal to every row of $\boldsymbol{A}$

Left Null Space

- Left null space is orthogonal to column space
- Each vector in left null space of a matrix is orthogonal to vectors in column space of the matrix



Big Picture of Linear Algebra



$$\mathbb{R}^n = \mathbf{N}(A) + \mathbf{C}(A^{\mathrm{T}})$$
$$\mathbf{N}(A) \perp \mathbf{C}(A^{\mathrm{T}})$$

$$\mathbb{R}^m = \mathbf{C}(A) + \mathbf{N}(A^{\mathrm{T}})$$
$$\boldsymbol{C}(A) \perp \boldsymbol{N}(A^{\mathrm{T}})$$

Part III

Projections

Linear Systems

Matrix Decomposition

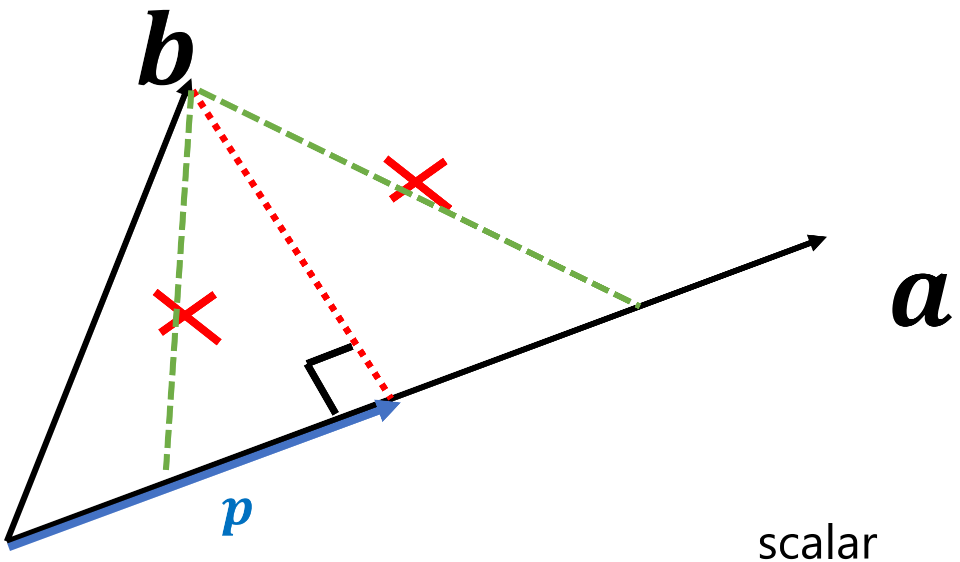
Projections

Projection onto a vector

Projection onto a space

$\boldsymbol{proj_a b}$: Projection of Vector $\boldsymbol{b}$ onto Vector $\boldsymbol{a}$

Vector in $span\{\boldsymbol{a}\}$ that is closest to vector $\boldsymbol{b}$



scalar

$$\boldsymbol{proj_a b} = \boldsymbol{p} = \hat{x}\ \boldsymbol{a}$$

$\boldsymbol{proj_a b}$: Projection of Vector $\boldsymbol{b}$ onto Vector $\boldsymbol{a}$

Dot product

$$\boldsymbol{e}.\boldsymbol{a} = 0$$

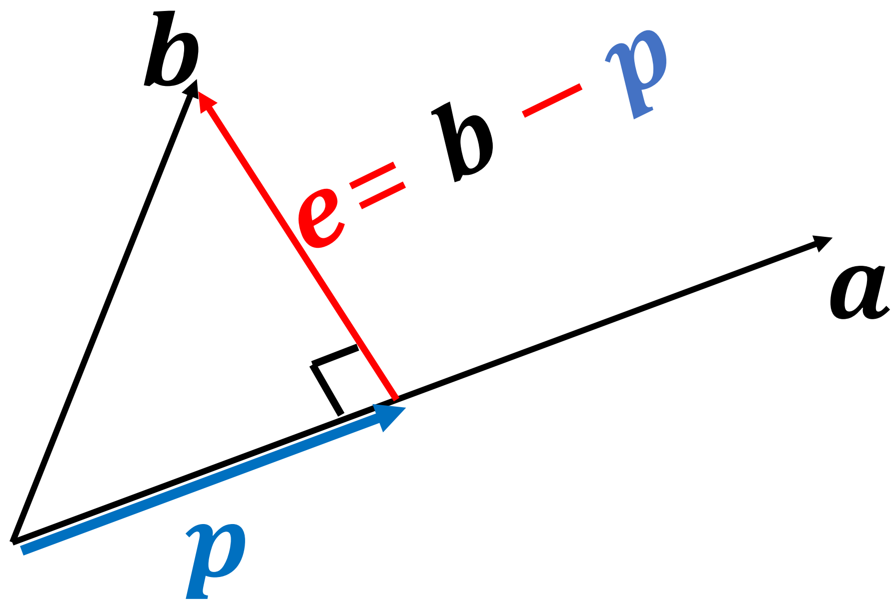
$$(\boldsymbol{b} - \boldsymbol{p}).\boldsymbol{a} = 0$$

$$(\boldsymbol{b} - \hat{x}\,\boldsymbol{a}).\boldsymbol{a} = 0$$

$$\hat{x}\,\boldsymbol{a}.\boldsymbol{a} = \boldsymbol{b}.\boldsymbol{a}$$

$$\hat{x} = \frac{\boldsymbol{b}.\boldsymbol{a}}{\boldsymbol{a}.\boldsymbol{a}}$$

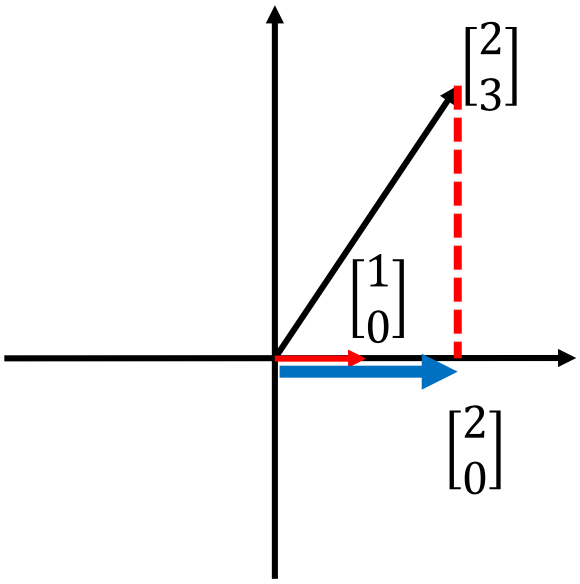
$$\Rightarrow \quad \boldsymbol{p} = \frac{\boldsymbol{b}.\boldsymbol{a}}{\boldsymbol{a}.\boldsymbol{a}}\boldsymbol{a}$$



$\boldsymbol{e}$ is orthogonal to $\boldsymbol{a}$

Projection onto a Vector: Example (1)

What is the projection of the vector $\boldsymbol{v}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ on to the vector $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$?



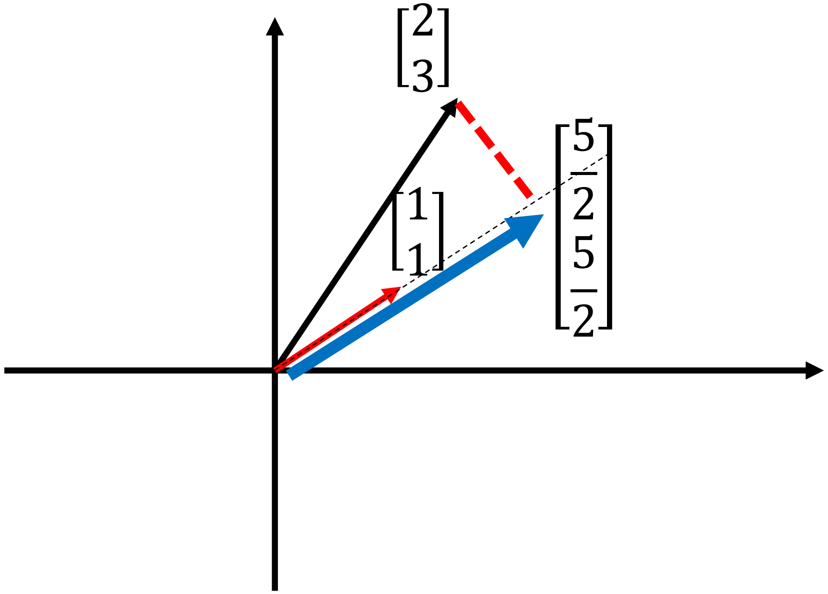
$$\boldsymbol{proj}_{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \frac{\begin{bmatrix} 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \frac{2}{1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad (\hat{x} = 2)$$

$$= \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Projection onto a Vector : Example (2)

What is the projection of the vector of vector $\boldsymbol{v}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ on to the vector $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$?



$$\boldsymbol{proj}_{\boldsymbol{v}_2}\boldsymbol{v}_1 = \frac{\boldsymbol{v}_1 \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2$$

$$\boldsymbol{proj}_{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \frac{\begin{bmatrix} 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

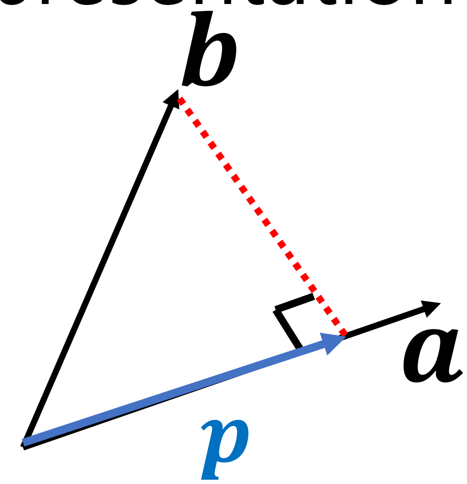
$$= \frac{5}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \left(\hat{x} = \frac{5}{2}\right)$$

$$= \begin{bmatrix} \frac{5}{2} \\ \frac{5}{2} \end{bmatrix}$$

https://en.wikibooks.org/wiki/Linear_Algebra/Orthogonal_Projection_Onto_a_Line

Projection onto a Vector: Alternate Representation

$$\boldsymbol{p} = \boldsymbol{a}(\boldsymbol{a}.\boldsymbol{a})^{-1}\ \boldsymbol{a}^T\boldsymbol{b} = \boldsymbol{P}\boldsymbol{b}$$



Proof: $\boldsymbol{proj_a b} = \hat{x}\boldsymbol{a}$

$= \boldsymbol{a}\,\widehat{x}$ because $\widehat{x}$ is a scalar

$= \boldsymbol{a}\,\frac{\boldsymbol{b}\cdot\boldsymbol{a}}{\boldsymbol{a}.\boldsymbol{a}}$ substitute $\widehat{x} = \frac{\boldsymbol{b}.\boldsymbol{a}}{\boldsymbol{a}.\boldsymbol{a}}$ is a scalar

$= \boldsymbol{a}(\boldsymbol{a}\,.\,\boldsymbol{a})^{-1}\ \boldsymbol{b}\cdot\boldsymbol{a}$ because $\boldsymbol{a}\,.\,\boldsymbol{a}$ is a scalar

$= \boldsymbol{a}(\boldsymbol{a}.\boldsymbol{a})^{-1}\ \boldsymbol{a}^T\boldsymbol{b}$ definition of inner product

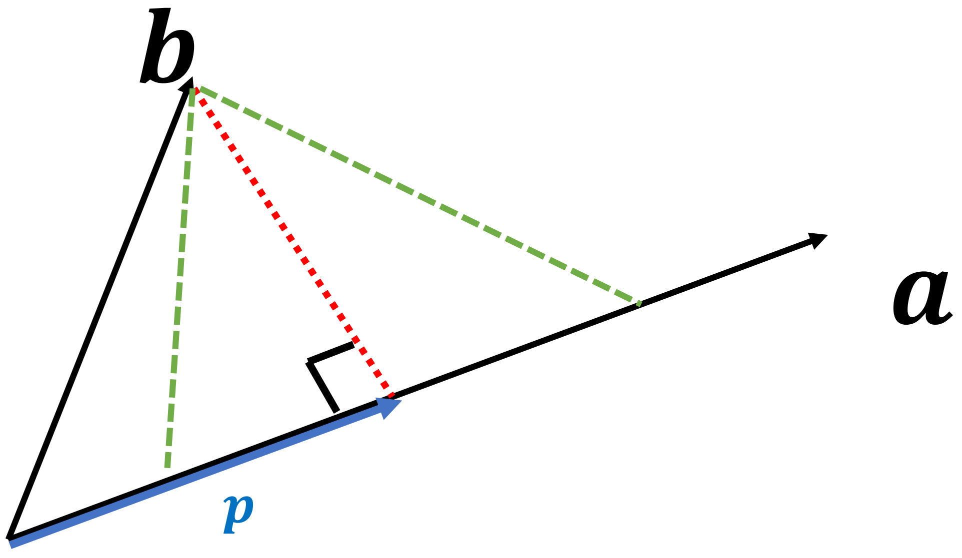
$\boldsymbol{p} = \boldsymbol{P}\boldsymbol{b}$

Projection onto a Vector: Projection Matrix

$\boldsymbol{p} = \boldsymbol{P}\boldsymbol{b}$, where $\boldsymbol{P} = \boldsymbol{a}(\boldsymbol{a}.\boldsymbol{a})^{-1}\ \boldsymbol{a}^T$

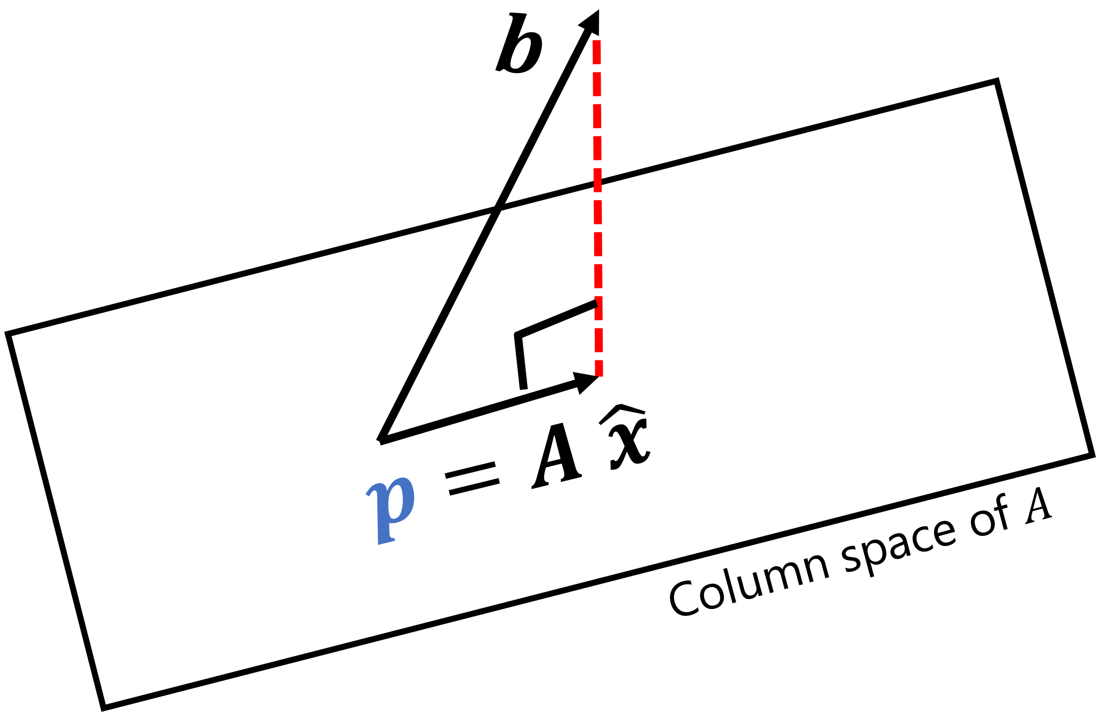
$\boldsymbol{P}$ is called projection matrix

$\boldsymbol{P}$ projects vector $\boldsymbol{b}$ into the $span\{\boldsymbol{a}\}$



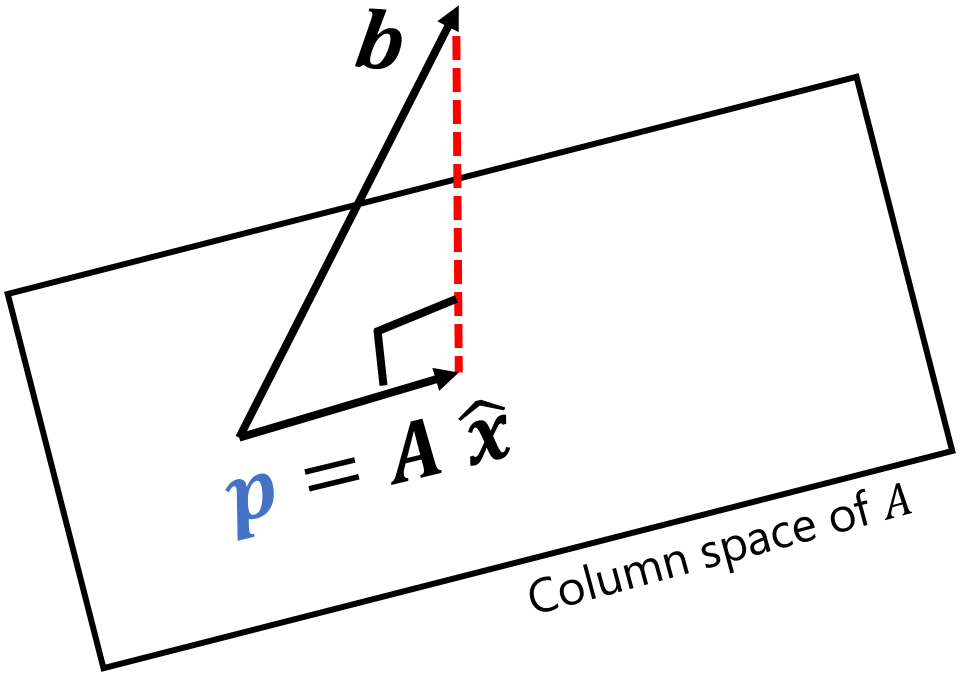
Projection onto a Space

- Project vector $\boldsymbol{b}$ onto the space spanned by the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_K$
- Or project vector $\boldsymbol{b}$ onto the column space of matrix $\boldsymbol{A} = [\boldsymbol{a}_1\ \boldsymbol{a}_2 \ldots\ \boldsymbol{a}_K]$



Projection onto a Space

$$\hat{\boldsymbol{x}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$$



$$\begin{aligned}\boldsymbol{p} &= \boldsymbol{A}\hat{\boldsymbol{x}} \\ &= \boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b} \\ &= \boldsymbol{P}\boldsymbol{b}\end{aligned}$$

- $\boldsymbol{P}\boldsymbol{b}$ is the vector in column space of $\boldsymbol{A}$ that is closest to vector $\boldsymbol{b}$

Projection Matrix

- $\boldsymbol{P} = \boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$
  - Projects a vector to the column space of $\boldsymbol{A}$

- Properties of projection matrix:
  - $\boldsymbol{P}^2 = \boldsymbol{P}$
  - $\boldsymbol{P}^T = \boldsymbol{P}$

Homework: Prove the properties of projection matrix $\boldsymbol{P}$

Linear System

System of linear equations

Solving a Linear System

$$x_1 + x_2 + x_3 = 1$$
$$2x_1 + 3x_2 + x_3 = 4$$
$$7x_1 - x_2 + x_3 = 6$$

$$\Rightarrow \quad \begin{array}{ccc} \boldsymbol{A} & \boldsymbol{x} & & \boldsymbol{b} \\ \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 7 & -1 & 1 \end{bmatrix} & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} & = & \begin{bmatrix} 1 \\ 4 \\ 6 \end{bmatrix} \end{array}$$

$$\Rightarrow \quad \boldsymbol{Ax} = \boldsymbol{b}$$

$\boldsymbol{A} \in \mathbb{R}^{\boldsymbol{m} \times \boldsymbol{n}}$ $\quad \boldsymbol{b} \in \mathbb{R}^{\boldsymbol{m}}$

$\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}$

m: Number of linear equations

n: Number of variables

Solving a Linear System

$$\boldsymbol{Ax} = \boldsymbol{b}$$

Attempt 1: Multiply with A^{-1}on both sides

$$\boldsymbol{A^{-1}Ax} = \boldsymbol{A^{-1}b}$$

$$\boldsymbol{x} = \boldsymbol{A^{-1}b}$$

What if $\boldsymbol{A}$ is not square?

Inverse exists only if matrix is square

What if $\boldsymbol{A}$ is square but singular?

Inverse does not exist for singular matrix

Linear System: Data Analytics Perspective

$$\boldsymbol{Ax} = \boldsymbol{b},\ \boldsymbol{A} \in \mathbb{R}^{\boldsymbol{m \times n}}$$

m: Number of measurements (data points)
n: Number of variables

- In practice, more measurements than variables
  - $m > n, i.e.,$ A is a rectangular matrix
  - No solution to the linear system
  - Instead try to find an approximate solution

Linear Systems: Approximate Solution

$$\boldsymbol{Ax} = \boldsymbol{b}$$

- Find an $\widehat{\boldsymbol{x}}$ such that $\boldsymbol{A}\widehat{\boldsymbol{x}} \approx \boldsymbol{b}$

- In other words:

$$\widehat{\boldsymbol{x}} = \arg\min_{x} (\|\boldsymbol{Ax} - \boldsymbol{b}\|_2)^2$$

known as Least Squares problem

Least Squares Problem

Attempt 2: $\boldsymbol{A\widehat{x} \approx b}$

Step 1: Multiply with A^T on both sides

$$A^T A \widehat{\boldsymbol{x}} = A^T \boldsymbol{b}$$

Square and symmetric matrix

Step 2: Multiply with $(A^T A)^{-1}$ on both sides

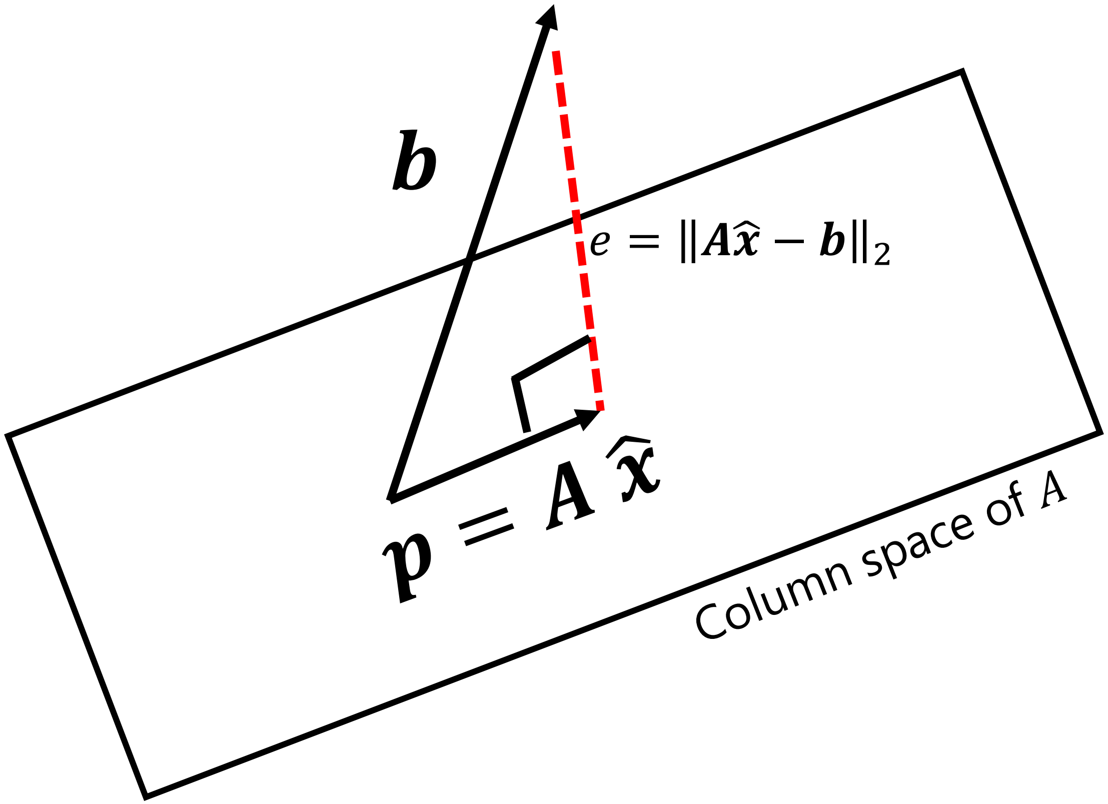
$$(A^T A)^{-1} A^T A \widehat{\boldsymbol{x}} = (A^T A)^{-1} A^T \boldsymbol{b}$$

$$\widehat{\boldsymbol{x}} = (A^T A)^{-1} A^T \boldsymbol{b}$$

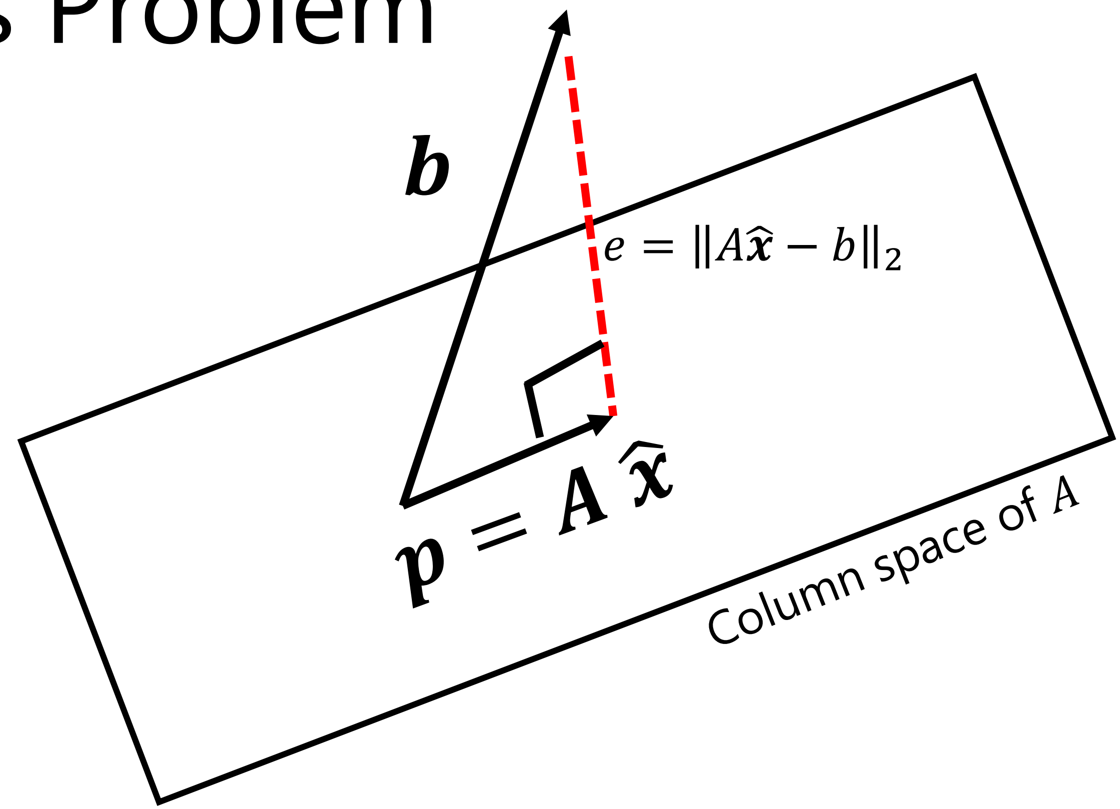
Recall: Projection Problem

$$\hat{\boldsymbol{x}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$$

$$\begin{aligned}\boldsymbol{p} &= \boldsymbol{A}\hat{\boldsymbol{x}} \\ &= \boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b} \\ &= \boldsymbol{P}\boldsymbol{b}\end{aligned}$$



Least Squares Problem



Known as residual error in ML

$$e = \|A\hat{\boldsymbol{x}} - \boldsymbol{b}\|_2 = \|A(A^TA)^{-1}A^T\boldsymbol{b} - \boldsymbol{b}\|_2$$

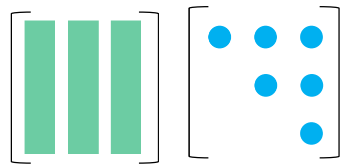
$$= \|\boldsymbol{P}\boldsymbol{b} - \boldsymbol{b}\|_2$$

Matrix Decomposition

Matrix Factorizations

- QR Factorization

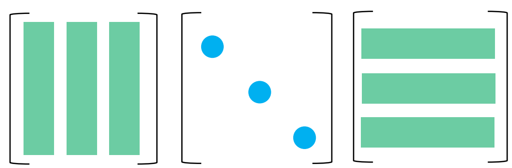
$$A = QR$$



QR decomposition as
Gram-Schmidt orthogonalization
Orthogonal $\boldsymbol{Q}$ and upper triangular $\boldsymbol{R}$

- Eigen Value Decomposition

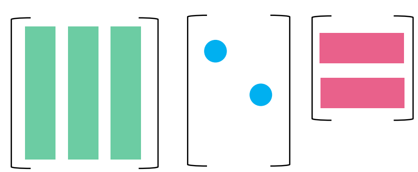
$$S = Q\Lambda Q^{\mathrm{T}}$$



Eigenvalue decomposition
of a symmetric matrix $\boldsymbol{S}$
Eigenvectors in $\boldsymbol{Q}$, eigenvalues in Λ

- Singular Value Decomposition

$$A = U\Sigma V^{\mathrm{T}}$$



Singular value decomposition
of all matrices $\boldsymbol{A}$
Singular values in Σ

Gram-Schmidt Orthogonalization

Factorization of a matrix ($\boldsymbol{A}$) into product of an orthogonal matrix ($\boldsymbol{Q}$) and upper triangular matrix ($\boldsymbol{R}$)

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$$

QR Factorization

$$\boldsymbol{A} = \boldsymbol{QR}$$

$$\begin{bmatrix} \\ \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \boldsymbol{q}_1 & \cdots & \boldsymbol{q}_n \\ \\ \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdot & r_{1n} \\ & r_{22} & \cdot & r_{2n} \\ & & \cdot & \cdot \\ & & & r_{nn} \end{bmatrix}$$

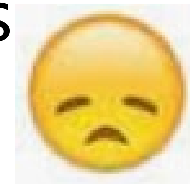
$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$: Linearly Independent Column Vectors

$\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$: Orthonormal Vectors

Why QR Factorization?

- Solving a set of linear equations

$$\boldsymbol{Ax = b}$$



Step 1: Factorize the matrix A

$$\boldsymbol{QRx = b}$$

Step 2: Multiply both sides by $\boldsymbol{Q^T}$

$$\boldsymbol{Q^T QRx = Q^T b}$$

$$\boldsymbol{Rx = Q^T b}$$

Why?

Why QR Factorization?

Without QR factorization

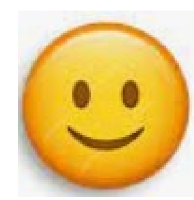
$A\boldsymbol{x} = \boldsymbol{b}$

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}\,x_3 &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}\,x_3 &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}\,x_3 &= b_3
\end{aligned}
$$

With QR factorization

$R\boldsymbol{x} = Q^T\boldsymbol{b}$

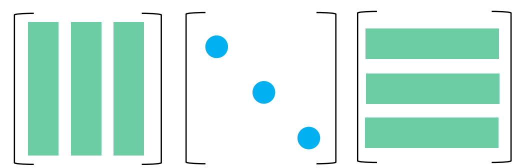
$$
\begin{aligned}
r_{11}x_1 + r_{12}x_2 + r_{13}\,x_3 &= c_1 \\
r_{22}x_2 + r_{23}\,x_3 &= c_2 \\
r_{33}\,x_3 &= c_3
\end{aligned}
$$



Matrix Decomposition

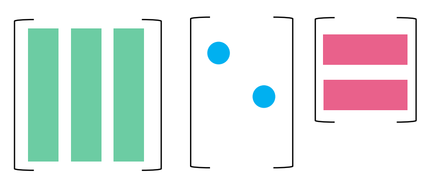
- Eigen Value Decomposition
  - aka Matrix Diagonalization
  - Matrix must be square (why?)

$$A = Q\Lambda Q^{-1}$$



- Singular Value Decomposition
  - Matrix can be rectangular

$$A = U\Sigma V^{T}$$



Square Matrix: Eigen Value Decomposition

- Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ be eigen vectors of a square matrix $\boldsymbol{A}$.

- $\boldsymbol{AQ} = \boldsymbol{A}\,[\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \boldsymbol{v}_3]$

$= [\boldsymbol{Av}_1 \quad \boldsymbol{Av}_2 \quad \boldsymbol{Av}_3]$ Matrix multiplication ($\boldsymbol{MM}3$)

$= [\lambda_1\boldsymbol{v}_1 \quad \lambda_2\boldsymbol{v}_2 \quad \lambda_3\boldsymbol{v}_3]$ Definition of eigen vectors

$$= [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \boldsymbol{v}_3]\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$ Using $\boldsymbol{Ax}\mathbf{2}$ **and** $\boldsymbol{MM}3$

$$= \boldsymbol{Q}\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

Square Matrix: Eigen Value Decomposition

- Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ be eigen vectors of a matrix A.

$$\boldsymbol{AQ} = \boldsymbol{Q}\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

$$\mathbf{A} = \boldsymbol{Q}\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}\boldsymbol{Q}^{-1}$$

Only if:

Matrix **Q** is invertible

Determinant of **Q** must be non-zero

Eigen vectors must be linearly independent

Square and Symmetric Matrix

- Eigen Value Decomposition
  - Eigen values are real
  - Eigen vectors are orthonormal
  - i.e., $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$,
  - i.e., $\boldsymbol{Q}^T = \boldsymbol{Q}^{-1}$

$$\boldsymbol{A} = \boldsymbol{Q}\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}\boldsymbol{Q}^T$$

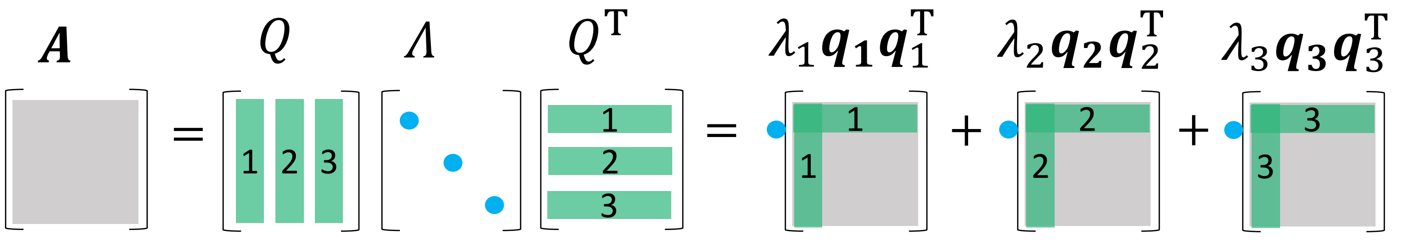
Applications of EVD

- Decompose a matrix into sum of rank-1 matrices

- Power of a matrix

EVD as Sum of Rank-1 Matrices

$$\boldsymbol{A} = \boldsymbol{Q}\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}\boldsymbol{Q}^T$$



EVD helps us decompose a matrix into a sum of rank-1 matrices

EVD in Power of a Matrix

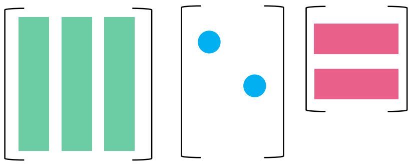
$$\boldsymbol{A} = \boldsymbol{Q} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \boldsymbol{Q}^T$$

$$\boldsymbol{A}^n = \boldsymbol{Q} \begin{bmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{bmatrix} \boldsymbol{Q}^T$$

Singular Value Decomposition

- SVD can be done for rectangular matrices also

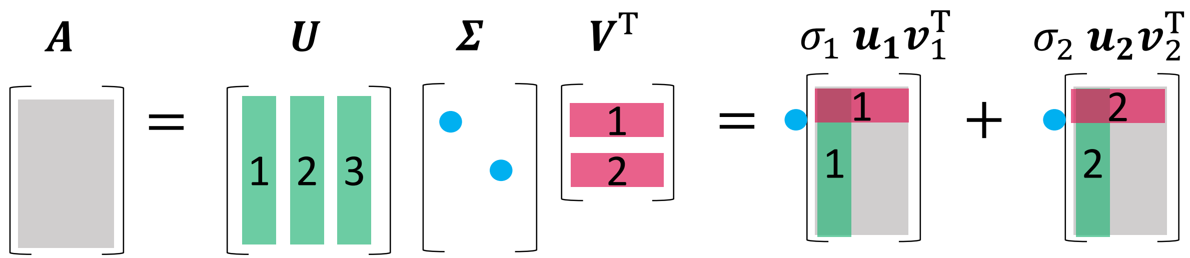
$$\boldsymbol{A} = \boldsymbol{U\Sigma V}^T$$



$\boldsymbol{U}$: Left singular matrix (orthogonal matrix)
$\boldsymbol{\Sigma}$: Diagonal matrix with singular values as diagonal elements
$\boldsymbol{V}$: Right singular matrix (orthogonal matrix)

SVD of a Matrix

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$$



Columns of $\boldsymbol{U}$ are the eigen vectors of $\boldsymbol{A}\boldsymbol{A}^T$
Columns of $\boldsymbol{V}$ are the eigen vectors of $\boldsymbol{A}^T\boldsymbol{A}$
Singular values are the eigen values of $\boldsymbol{A}\boldsymbol{A}^T$ or $\boldsymbol{A}^T\boldsymbol{A}$

SVD of a Matrix

$$\boldsymbol{AV} = \boldsymbol{U\Sigma V}^T\boldsymbol{V}$$
$$= \boldsymbol{U\Sigma}$$

$$A\begin{bmatrix} & & \\ \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_r \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r \\ & & \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

$$= [\ \sigma_1\boldsymbol{u}_1 \quad \dots \quad \sigma_r\boldsymbol{u}_r]$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$$

r = Rank of $\boldsymbol{A}$