HCAI5TML01 – Mathematics of Learning.
Week – 1: Lecture – 02
A Refresher on the Mathematics Behind Machine Learning.
Linear Algebra – Towards Fundamental Theory of Linear Algebra.

Siman Giri

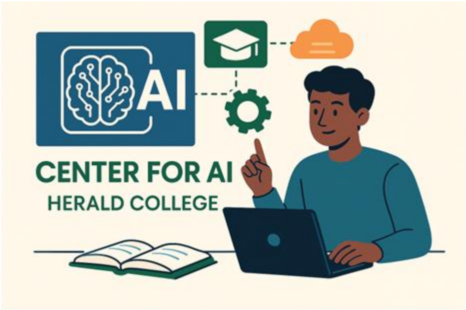


image generated via copilot.

HERALD COLLEGE | CENTER FOR AI.

Exercise

Problem 1: Span

Determine whether the vector $\mathbf{v} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ is in the span of the vectors $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Problem 2: Basis

Do the vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ form a basis of $\mathbb{R}^3$?

1. The Fundamental Subspace of a Matrix.

1.1 Matrix as a Function: Linear Transformation.

- For a matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$,
  - we can view **this matrix as a function** that **maps** vectors from $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$.
- This mapping is implemented by **matrix – vector multiplication** i.e. $\mathbf{Ax = b}$.
  - Here, A **vector** $\mathbf{x} \in \mathbb{R}^n$ is **mapped** to **vector** $\mathbf{b} \in \mathbb{R}^m$.
- Stated as Linear Transformation:
  - We can define a function $\mathbf{T}: \mathbb{R}^n \to \mathbb{R}^m$ as:
    - $\mathbf{T(x) \coloneqq Ax}$

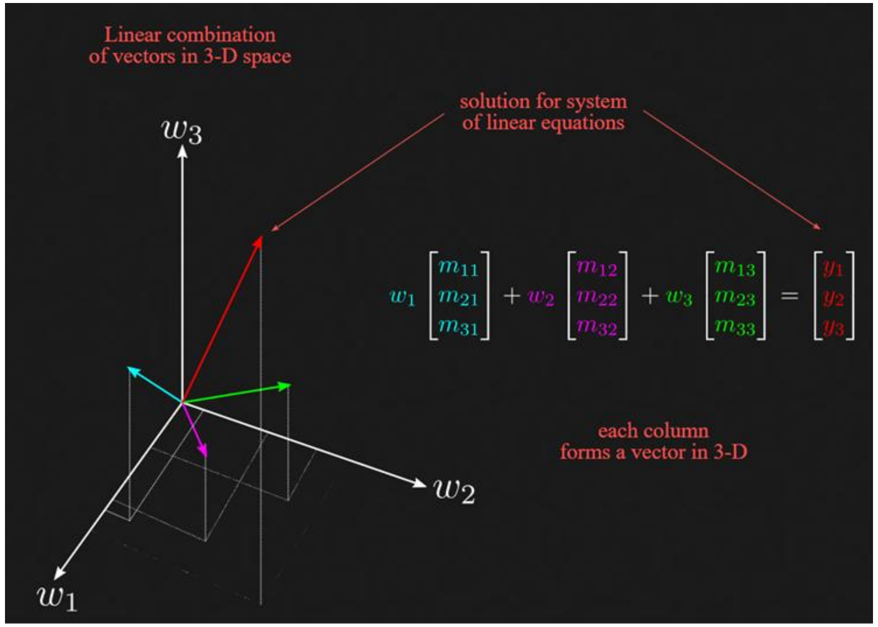


Fig: Solution of Ax {image: pablo (©catbug88)}.

- **What is the Range of a Matrix?**
  - For matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$, The **range of A** (aka **image of the linear transformation A**),
    - is the set of all **vectors in** $\mathbb{R}^m$ that can be written **as Ax for some** $\mathbf{x} \in \mathbb{R}^n$.
      - $\mathbf{Range(A) = \{b \in \mathbb{R}^m | b = Ax \text{ for some } x \in \mathbb{R}^n\}}$.

1.1 Matrix subspaces

- Let's **recall the definition of a subspace** in the **context of vectors**:
  - **Contains the zero vector, $0 \in S$**
  - **Closure under multiplication, $\forall \alpha \in \mathbb{R} \rightarrow \alpha \times s_i \in S$**
  - **Closure under addition, $\forall s_i \in S \rightarrow s_1 + s_2 \in S$**
- Since matrices are collections of vectors (as rows or columns), we can explore all the subspaces formed by their structure.
  - Thus, now we can ask what are all possible subspaces that can be "covered" by a collection of vectors in a matrix.
- There are four fundamental subspaces that can be "covered" by a matrix of valid vectors, Hence called :
  - "**The Four Fundamental Subspaces**"

- **What is the Range of a Matrix?**
  - For matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{m} \times \mathbf{n}}$, The **range of A** (aka **image of the linear transformation A**),
    - is the set of all **vectors in $\mathbb{R}^{\mathbf{m}}$** that can be written **as $\mathbf{Ax}$ for some $\mathbf{x} \in \mathbb{R}^{\mathbf{n}}$**.
      - **$\mathrm{Range}(\mathbf{A}) = \{\mathbf{b} \in \mathbb{R}^{\mathbf{m}} | \mathbf{b} = \mathbf{Ax}$ for some $\mathbf{x} \in \mathbb{R}^{\mathbf{n}}\}$** .

Subspaces of matrix A

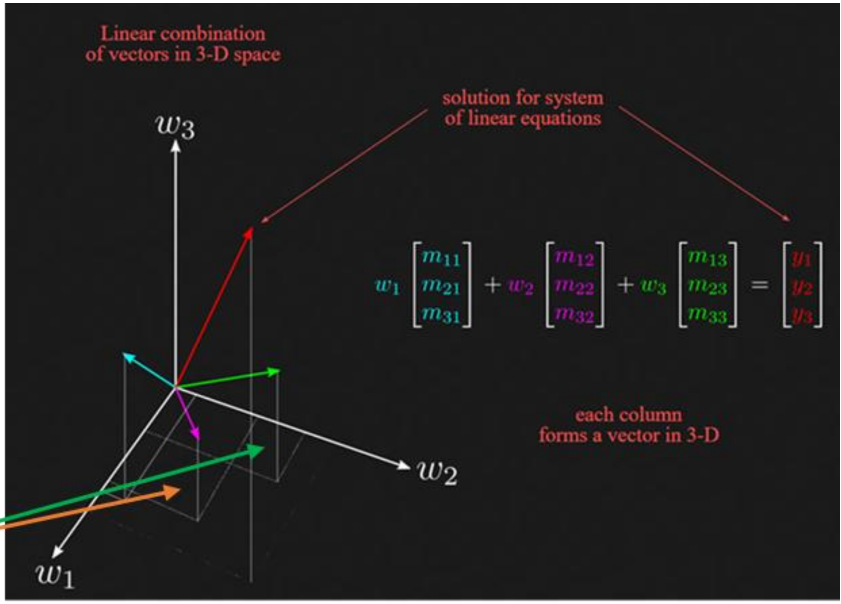


Fig: Solution of Ax {image: pablo (©catbug88)}.

1.2 The Four Fundamental Subspaces.

Let $\mathbf{A} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$. Then:

Subspace	Symbol	Defined in:	Dimension	Description
Column Space	$\mathbf{Col}(\mathbf{A})$	$\mathbb{R}^{m}$	$\mathbf{rank}(\mathbf{A})$	All vectors $\mathbf{b}$ for which $\mathbf{Ax} = \mathbf{b}$ is solvable.
Null Space	$\mathbf{Null}(\mathbf{A})$	$\mathbb{R}^{n}$	$\mathbf{rank} - \mathbf{rank}(\mathbf{A})$	All solutions to $\mathbf{Ax} = \mathbf{0}$.
Row Space	$\mathbf{Row}(\mathbf{A})$	$\mathbb{R}^{n}$	$\mathbf{rank}(\mathbf{A})$	The span of rows of $\mathbf{A}$ (or columns of $\mathbf{A}^{\mathbf{T}}$).
Left Null Space	$\mathbf{Null}(\mathbf{A}^{\mathbf{T}})$	$\mathbb{R}^{m}$	$\mathbf{m} - \mathbf{rank}(\mathbf{A})$	All $\mathbf{y} \in \mathbb{R}^{\mathbf{m}}$ such that $\mathbf{A}^{\mathbf{T}}\mathbf{y} = \mathbf{0}$.

1.2.1 Column Spaces.

- The **column space** of a matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$, denoted as $\mathbf{C(A)}$,
  - is the set of all **linear combinations** of its **column vectors.**
    - For any Matrix A , the column space of A, is the vector space that spans the column vectors of A.
- **Formal Definition:**
  - Let $\mathbf{A} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$. The column space of A is: $\mathbf{C(A) = \{b \in \mathbb{R}^m | b = Ax \text{ for some } x \in \mathbb{R}^n\}}$
- **Interpretation:**
  - Each **column of A** is a vector in $\mathbb{R}^{\mathbf{m}}$.
  - The **column space** is **the span of these vectors**.
  - Any **output b** $\in \mathbf{C(A)}$ is formed by multiplying with some **vector x** $\in \mathbb{R}^{\mathbf{n}}$.
  - Determines the **range** of the **linear transformation A.**
    - $\mathbf{Range(A) = \{b \in \mathbb{R}^m | b = Ax \text{ for some } x \in \mathbb{R}^n\} = C(A)}$.

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

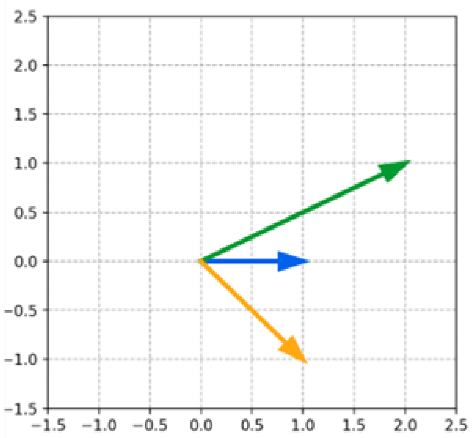


Fig: Matrix as a column vector.

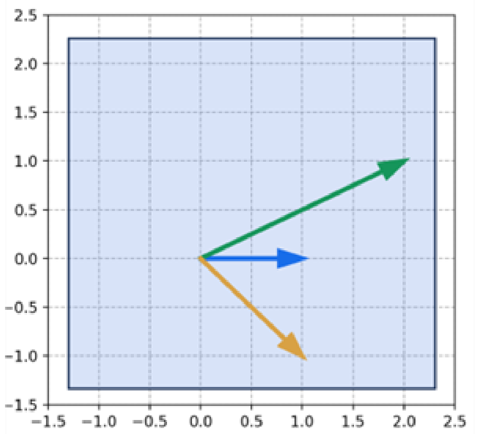


Fig: Column space.

1.2.2 Row Spaces

- The row space of a matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{m \times n}}$, denoted as $\mathcal{R}(\mathbf{A})$, is the set of all linear combinations of its row vectors.
- **Formal Definition:**
  - Let $\mathbf{A} \in \mathbb{R}^{\mathbf{m \times n}}$. Then the row space of A is:
    - $\mathcal{R}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^{\mathbf{n}} \mid \mathbf{y} = \boldsymbol{\alpha}_1\mathbf{r}_1 + \boldsymbol{\alpha}_2\mathbf{r}_2 + \cdots + \boldsymbol{\alpha}_m\mathbf{r}_m \textbf{ for some } \boldsymbol{\alpha}_i \in \mathbb{R}\}$ or can also be written as:
      - $\mathcal{R}(\mathrm{A}) =$ **span of the row vectors of A.**
- **Interpretation:**
  - Each row of A is a vector in $\mathbb{R}^n$
  - The row space is a subspace of $\mathbb{R}^n$
  - It contains all possible linear combinations of the rows.
- **Key properties:**
  - Dimension of **row space = rank(A).**
  - The row space of A is the column space of $\mathbf{A^T}$: $\mathcal{R}(\mathbf{A}) = \mathbf{C}(\mathbf{A^T})$
  - Row space **captures constraints** on the **input x** in solving $\mathbf{Ax = b}$.

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

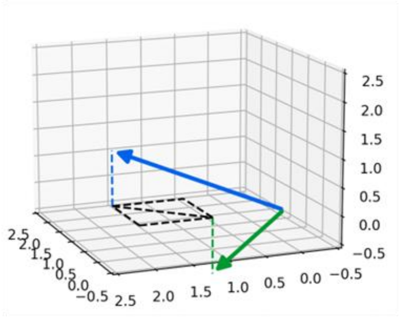


Fig: Matrix as a row vector.

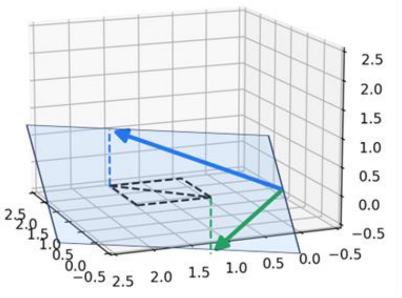


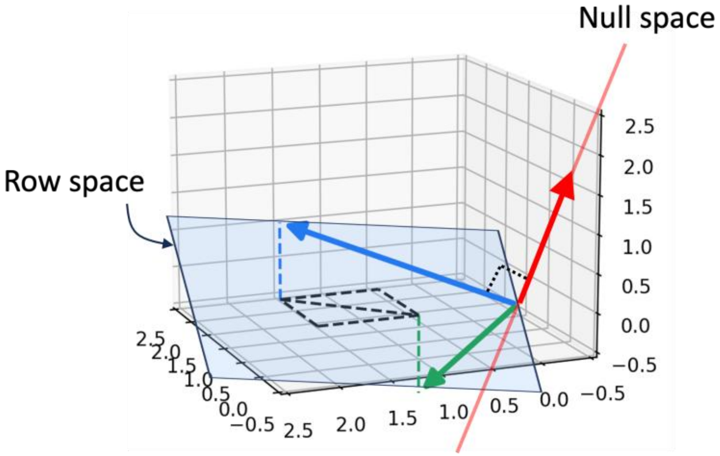
Fig: Row space.

Constraint in Linear Algebra:

- A **constraint** is a rule or condition that a solution must satisfy.
  - In the context of solving a system of equations like: $Ax = b$
  - Each row of the matrix A represents a linear equation,
    - and each of these equation is constraint on the unknown vector x.
- Example:
  - Let $\mathbf{A} = \begin{bmatrix} \mathbf{1} & \mathbf{2} \\ \mathbf{3} & \mathbf{1} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} \mathbf{5} \\ \mathbf{7} \end{bmatrix}$
  - This gives two constraints:
    - $\begin{matrix} \mathbf{x_1 + 2x_2 = 5} \\ \mathbf{3x_1 + x_2 = 7} \end{matrix}$ so, vector x must satisfy both equations at same time,
      - each equation constraints what the value or solution of $\mathbf{x} \in \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix}$ can be.
- Interpretation:
- The solution x must lie in space of inputs that satisfy all linear constraints defined by the rows.
  - Must lie in the subspace of Row vectors i.e. Row space.
- If multiple rows are linearly dependent some constraints are redundant i.e., they give the same constraint, just scaled not adding any new information.

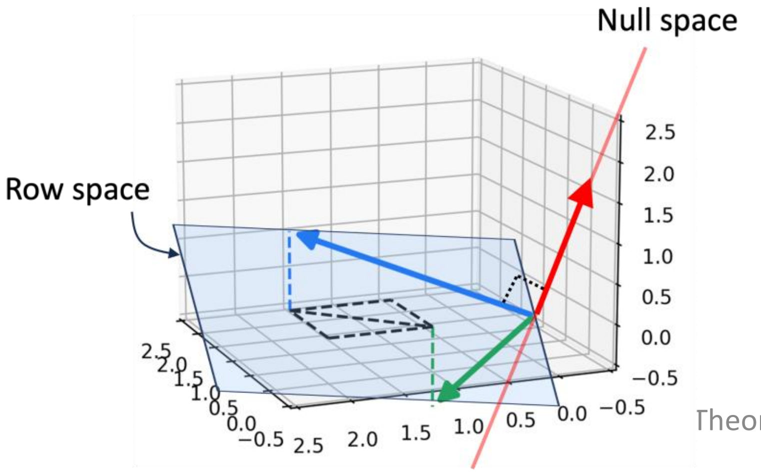
1.2.3 Null Space.

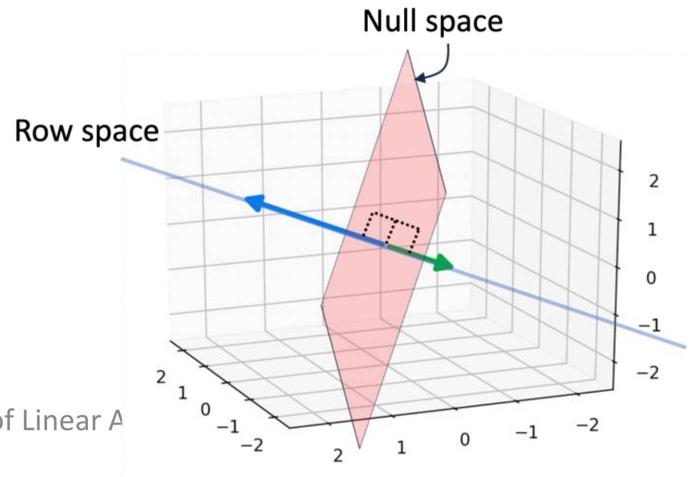
- Let $\mathbf{A} \in \mathbb{R}^{\mathbf{m\times n}}$.
  - The null space of A, denotes as $\mathcal{N}(\mathbf{A})$, is the **set of all vectors x** $\in \mathbb{R}^{\mathbf{n}}$ that get **mapped to the zero** vector when multiplied by A:
    - $\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^{\mathbf{n}} | \mathbf{Ax} = \mathbf{0}\}$
  - So, It is the set of all input vectors that get mapped to the zero vector in output.
  - The entire direction of x is lost or flattened by the transformation.
- **Interpretation:**
  - Each vector in the null space is an **input** $\mathbf{x} \in \mathbb{R}^{\mathbf{n}}$ that gets sent to $\mathbf{0} \in \mathbb{R}^{\mathbf{m}}$
  - The null space lives in $\mathbb{R}^n$ (input space)
  - If A is full rank, **null space**$\{\mathbf{0}\}$
  - If A is rank – deficient, null space contains infinitely many directions.
  - The dimension of the null space is called the nullity: $\mathbf{nullity(A) = n - rank(A)}$
    - This tells you the **number of independent directions in input space** that are **flattened to zero** by A.
      - nullity tells you how many input dimensions are invisible in the output



1.2.4 Left Null Space – Null Space of Transpose.

- Let $\mathbf{A} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$.
  - The left null space of A , also known as the **null space of $\mathbf{A}^{\mathbf{T}}$** is:
    - $\mathcal{N}(\mathbf{A}^{\mathbf{T}}) = \{\, \mathbf{y} \in \mathbb{R}^{\mathbf{n}} | \mathbf{A}^{\mathbf{T}}\mathbf{y} = \mathbf{0}\}$
- Interpretation:
  - It is the set of all vectors in $\mathbb{R}^m$ that are orthogonal to rows of A.
  - In layman terms:
    - What vectors, when dotted with each row of A, give zero?
  - So,
    - **While $\mathcal{N}(\mathbf{A}) \subset \mathbb{R}^{\mathbf{n}}$ (input directions killed by A)**
    - **$\mathcal{N}(\mathbf{A}^{\mathbf{T}}) \subset \mathbb{R}^{\mathbf{m}}$ (output directions orthogonal to the row space of A)**





Fundamental Subspaces of Matrix.

{Example Walkthrough with a Matrix.}

Example.

- Let's take:
  - $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3}$
  - Compute all Four Subspaces:
    1. **Step 1: Row reduced to Echelon Form:**
       - Apply row operations:
         - $\begin{matrix} R_2 \leftarrow R_2 - 2R_1 \\ R_3 \leftarrow R_3 - R_1 \end{matrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 1 & -2 \end{bmatrix}$
         - $R_3 \leftarrow -1 \cdot R_3$ and swap $\Rightarrow A_{ref} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$
       - Identify the rank of Matrix:
         - $\text{rank}(A) = 2$
           - As there are **two pivots** in $\mathbf{A}_{\mathbf{ref}}$ matrix.
           - What are Pivots?

Understanding Pivot and Rank.

- **Pivots** are the first non-zero entries in each row of a matrix after it has been transformed into **row echelon form (REF)** or **reduced row echelon form (RREF)**.
- **Key Properties:**
  - **Location:**
    - Each pivot must be the **leftmost non-zero element** in its row.
    - Pivots shift to the right as you move down the rows.
  - **Structure**:
    - In **REF**: Pivots can be any non-zero number (often normalized to 1).
    - In **RREF**: Pivots are **1**, and all entries above/below them are **0.**
  - **Uniqueness**:
    - Each row and column can contain **at most one pivot**.
    - The number of **pivots = rank of the matrix**.
    - In the example above: **2 pivots → rank = 2.**
  - **Singular (Rank-Deficient) Matrix**:
    - At least one row lacks a pivot

Pivot.

$$A_{\text{ref}} = \begin{bmatrix} \boxed{1} & 2 & 3 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Good to Know: REF vs RREF

Property	REF	RREF
Leading Entry	First non-zero entry in a row (pivot) can be any non-zero number.	Pivots must be 1.
Pivot Columns	Entries below pivots are 0.	Entries above and below pivots are 0.
Uniqueness	Not unique (multiple REFs possible).	Unique for a given matrix.

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

REF or RREF.

Compute Column and Row Space.

- **Column Space:**
  - Use original columns (1^{st} and 2^{nd}):
    - $\mathbf{Col(A) = span}\left\{\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}2\\4\\1\end{bmatrix}\right\} \subset \mathbb{R}^3$
    - $\mathbf{dim\big(Col(A)\big) = 2}$
    - "Pivot columns in RREF → corresponding columns in original A."
- **Row space:**
  - Take non – zeros from row reduced form:
    - $\mathcal{R}\mathbf{(A) = span\{[1 \quad 2 \quad 3], [0 \quad 1 \quad 2]\}} \subset \mathbb{R}^3$
    - "Non-zero rows in RREF (not original matrix!) define the row space."

Null and Left Null Space

- **Null Space:**
  - Solve $\mathbf{Ax} = \mathbf{0}$
    - Form REF:
      - $\begin{array}{r} \mathbf{x_1 + 2x_2 + 3x_3 = 0} \\ \mathbf{x_2 + 2x_3 = 0} \end{array} \Rightarrow \mathbf{x_2 = -2x_3,\ x_1 = x_3}$
  - General Solution:
    - $\mathbf{x = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \Rightarrow Null(A) = span\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right\} \subset \mathbb{R}^3}$
- **Left Null Space:**
  - $\mathbf{A^T = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 1 \\ 3 & 6 & 1 \end{bmatrix}}$
  - Solve $\mathbf{A^T y = 0}$**,** i.e.,
    - $\begin{array}{r} \mathbf{y_1 + 2y_2 + y_3 = 0} \\ \mathbf{2y_1 + 4y_2 + y_3 = 0} \\ \mathbf{3y_1 + 6y_2 + y_3 = 0} \end{array} \Rightarrow \mathbf{solution: y = y_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \Rightarrow Null\left(A^T\right) = span\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}}$

Summary

Subspace	Space it lies in	Dimension	Spanned by
Column Space	$\mathbb{R}^3$	2	First and Second Column of A.
Null Space	$\mathbb{R}^3$	1	$[1 \quad -2 \quad 1]^T$
Row Space	$\mathbb{R}^3$	2	Two non – zero rows from REF of A
Left Null Space	$\mathbb{R}^3$	1	$[-1 \quad 0 \quad 1]^T$

What does it mean for Machine Learning?

Exercise.

Problem 3: Null Space

Find the null space of the matrix

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

1.3 Revisiting: Rank.

- The rank of a **matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{m \times n}}$** is defined as:
    - $\text{rank}(A) = \text{dimension of the column space of A.}$
    - That is, it is the maximum number of linearly independent columns (or rows) in the matrix.
- Equivalently:
    - $\textbf{rank}(\mathbf{A}) = \textbf{number of pivot columns in row echelon form of A.}$
- **Column rank = Row rank**; Even though they come from different spaces.
    - Column space is in $\mathbb{R}^m$ and
    - Row space is in $\mathbb{R}^n$
- The **number of linearly independent columns = number of linearly independent rows**, always.

1.3.1 Rank: Number of Useful Directions.

- In linear algebra and machine learning,
  - **rank** tells us **how much of the input space survives** after applying a **linear transformation (represented by a matrix).**
- **Formal Statement:**
  - The **rank** of a **matrix A** $\in \mathbb{R}^{\mathbf{m \times n}}$ is the **dimension of the column space** (or row space).
    - It tells you **how many linearly independent directions** (basis vectors) are preserved or **not collapsed** into zero.
- Interpretation:
  - Think of a matrix as a **machine** that transforms input vectors $x \in \mathbb{R}^n$ into outputs $Ax \in \mathbb{R}^m$.
  - Some directions get **flattened** or **lost** in the process (they become zero or redundant).
  - The **rank** tells you how many directions **don't get flattened** — i.e., the **useful directions**.

Example:

- **Full Rank (rank = n)**:
  - No information lost.
  - Transformation is **invertible** (if square).
  - All directions in input space are useful.
  - E.g., Identity matrix.

- **Rank 1 matrix** in $\mathbb{R}^{3\times3}$:
  - All input vectors get mapped to a **line**.
  - Only **one direction** is preserved.
  - The rest collapse (projected onto that line).
  - So: only 1 "useful direction."

1.3.2 Rank and Machine Learning Example.

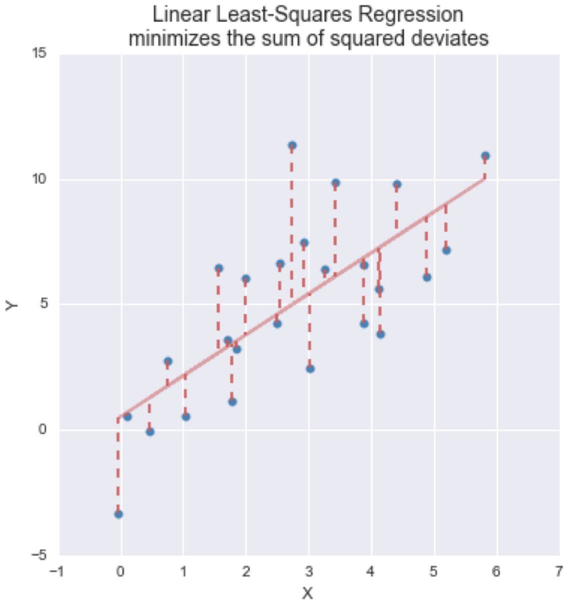
- Suppose you have a **feature matrix $\mathbf{X} \in \mathbb{R}^{\mathbf{n \times p}}$**, but many **features are linear combinations** of each other (collinearity).
  - The **$\mathbf{rank(X) < p}$**, and you **are not learning from p truly independent directions**.
  - Techniques like PCA, regularization, and SVD help isolate these useful directions.

2. Col Space and Machine Learning.

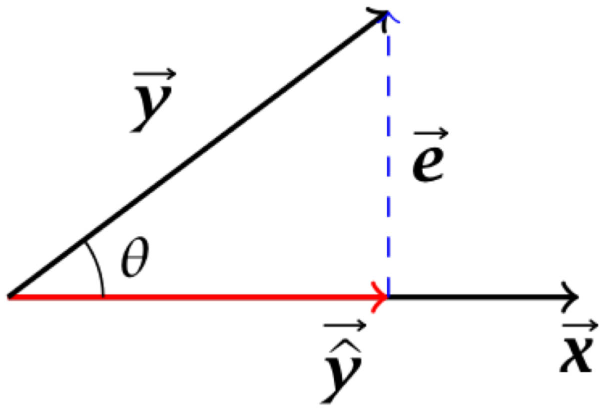
2.1 What Fundamental Subspace tells about data?

- For a **feature matrix**:
  - $\mathbf{X} \in \mathbb{R}^{\mathbf{n} \times \mathbf{d}}$ **(with n samples, and d features),**
    - the four fundamental subspaces reveal critical insights about your data and model.
- In upcoming slides, we will explore each subspace with a context of **linear regression problem**.

A)



B)

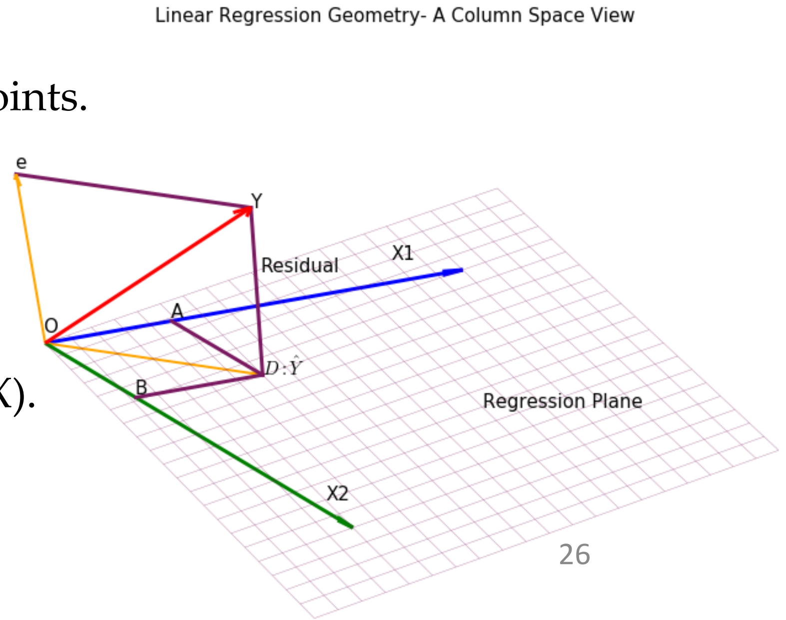


2.2 Linear Regression and the Col Space: A Geometric Intuition.

- **Linear Regression as a Linear System:**
  - In linear regression, we model the relationship: $y = X\beta + \epsilon$,
    - Where:
      - $X \in \mathbb{R}^{n \times d} \rightarrow$ is the feature matrix (each row is a simple, each column feature)
        $y \in \mathbb{R}^{n} \rightarrow$ is the target vector
        $\beta \in \mathbb{R}^{d} \rightarrow$ is the coefficient vector to find,
        $\epsilon \in \mathbb{R}^{n} \rightarrow$ is the error/residual
- **The goal is to solve: $X\beta \approx y$.**

2.2 Linear Regression and the Col Space: A Geometric Intuition.

- **Column Space: The "Reachable" Outputs:**
  - The column space of X (Col(X)) is the span of the feature vectors (columns of X).
  - It contains all possible linear combinations of the features, i.e. all possible predictions Xβ.
- **Key insights:**
  - If $\mathbf{y} \in \mathbf{Col(X)}$, there exists a perfect solution β such that $\mathbf{X\beta} = \mathbf{y}$.
  - If $\mathbf{y} \notin \mathbf{Col(X)}$, not exact solution exists, and we seek the best approximation (least squares).
- **Geometric Interpretations:**
  - Perfect Fit (Rare):
    - If y lies exactly in Col(X) , the regression line/passes through all data points.
    - This is almost impossible in real world scenario.
  - Approximate fit:
    - If y in not in Col(X), the best we can do is
      - project y onto Col(X), minimizing $\|y - X\beta\|^2$ called residual.
        - The residual is orthogonal to Col(X) i.e. it lies in Left NULL(X).



To summarize:

Subspace	Dimension (if $A \in \mathbb{R}^{m \times n}$)	ML meaning
Column space	rank(A)	All reachable predictions (range of model).
Null Space	n − rank(A)	Input directions that have no effect (feature redundancy)
Row Space	rank(A)	What constraints are imposed on inputs (in dual space)
Left Null Space	m − rank(A)	Residual errors in least squares, inconsistency.

Exercise

Problem 4: Fundamental Subspaces

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \\ 5 & 10 \end{bmatrix}$. Find the rank, nullity, and describe the fundamental subspaces.

3. Getting Started with Matrix Decompositions.

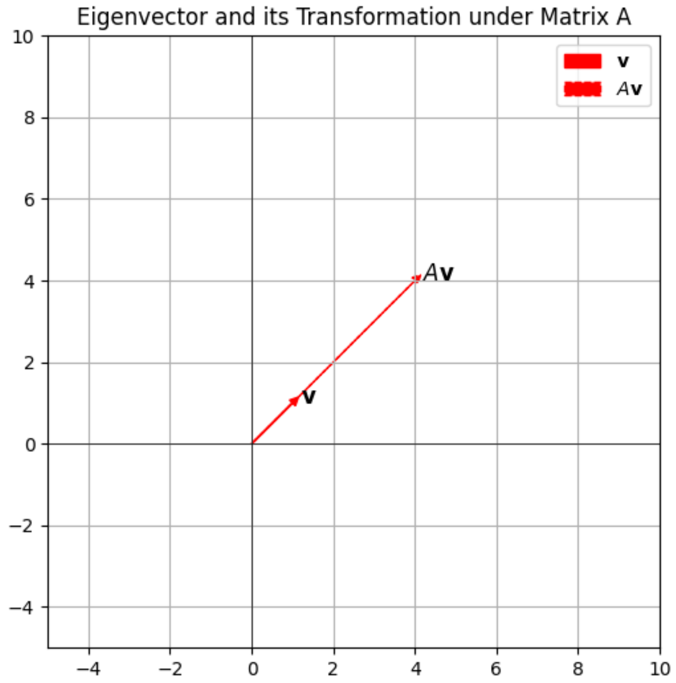
{Eigen Value Problem aka eigen – value Decompositions.}

3.1 Eigen Vector and Eigen Value.

- An **eigenvector** of a **square matrix A** is a **non-zero vector v** that, when **multiplied by A,** results in a scalar multiple of itself.
  - In other words, it is a vector that does not change direction when the linear transformation represented by A is applied to it.
    - Instead, it only gets scaled by a certain factor, called the **eigenvalue**.
- Mathematically, for any **Matrix – vector** pair if following holds:
  - $\boldsymbol{Av = \lambda v}$
    - then the **vector $\boldsymbol{v}$** is called eigen vector and the **scaling factor $\boldsymbol{\lambda}$** is called eigen value.
- **Key points about eigen vectors:**
  - **Non-zero**: Eigenvectors are always non-zero vectors, $\mathbf{i.e. v \neq 0}$.
  - **Scaling:** The transformation A simply scales the eigenvector by the **eigenvalue $\boldsymbol{\lambda}$**;
    - it does not change the **vector's direction**.
  - **Multiple eigenvectors**: For each eigenvalue, there can be infinitely many eigenvectors, all scalar multiples of each other. They form a subspace (called the eigenspace) corresponding to that eigenvalue.

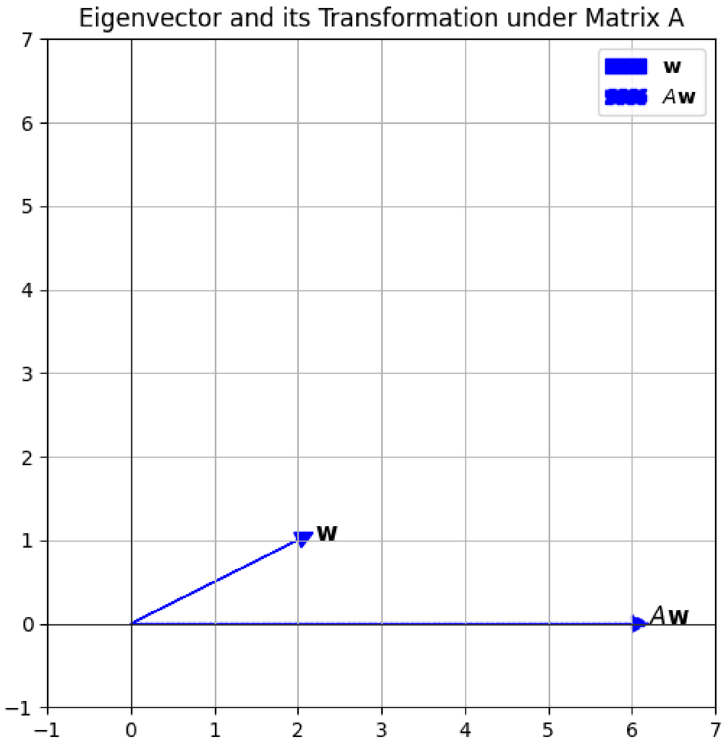
4.1.1 Identify the Eigen Vector.

- **Consider a matrix**
  - $\mathbf{A} = \begin{bmatrix} \mathbf{2} & \mathbf{2} \\ \mathbf{-4} & \mathbf{8} \end{bmatrix}$ and
    - vectors $\mathbf{v} = \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}$; $\mathbf{w} = \begin{bmatrix} \mathbf{2} \\ \mathbf{1} \end{bmatrix}$.
- **Which are Eigen vectors?**
  - For $\mathbf{v} \rightarrow$ **we check if v is an eigenvector by calculating Av**:
    - $\mathbf{Av} = \begin{bmatrix} \mathbf{2} & \mathbf{2} \\ \mathbf{-4} & \mathbf{8} \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix} = \begin{bmatrix} \mathbf{4} \\ \mathbf{4} \end{bmatrix} = \mathbf{4} \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix} \Rightarrow \mathbf{4v.}$
    - So, $\mathbf{v}$ is an **eigen vector** with **eigenvalue** $\boldsymbol{\lambda} = \mathbf{4}$.
- What about w?



3.1.1 Identify the Eigen Vector.

- **Consider a matrix**
  - $\mathbf{A} = \begin{bmatrix} 2 & 2 \\ -4 & 8 \end{bmatrix}$ and
    - vectors $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$; $\mathbf{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.
- **Which are Eigen vectors?**
  - For w → **we check if *w* is an eigenvector by calculating Aw**:
    - $\mathbf{Aw} = \begin{bmatrix} 2 & 2 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \neq \lambda w.$
    - So, **w** is not an **eigen vector there does not exist a**
      - **scalar λ** under which $\mathbf{Aw} = \lambda \mathbf{w}$ holds true.



3.2 Eigen Value Problem.

- The **eigenvalue problem** is a fundamental concept in linear algebra and plays a critical role in various fields such as machine learning, physics, and computer science.
- It involves **finding scalar values** (called **eigenvalues**) and corresponding **non-zero vectors** (called **eigenvectors**) for a given **square matrix**.
  - Mathematically, Given a **square matrix A**, the eigenvalue problem is to find **scalars λ** and **eigen vector v** that satisfy the following equation.
    - $\mathbf{Av} = \lambda\mathbf{v}$.

3.3 Steps to solve the Eigenvalue Problem.

- Write the characteristic equation:
  - To find the eigenvalues, we rewrite the equation as:
    - $(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I})\mathbf{v} = \mathbf{0}$ {called characterstic equation}
    - Where:
      - **I** is the **identity matrix** of the **same size as** A,
      - $\boldsymbol{\lambda} \rightarrow$ **eigen values**.
      - $\mathbf{v} \rightarrow$ **eigen vector**.
      - *Cautions: the matrix $A - \lambda I$ must be singular i.e. $det(A - \lambda I) = 0$.*
- Compute the characteristic polynomial:
  - Solve
    - $\mathbf{det(A - \lambda I) = 0}$,
  - which gives a **polynomial equation in $\boldsymbol{\lambda}$** which is called characteristic polynomial.
- Solve the characteristic polynomial:
  - Solve the polynomial equation to find the eigenvalues $\boldsymbol{\lambda_1, \lambda_2, \ldots \lambda_n}$.
- Find the eigen vectors:
  - For each **eigen value $\boldsymbol{\lambda_i}$**,
    - substitute it back into the equation $(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I})\mathbf{v} = \mathbf{0}$ and solve for the **eigenvector $\boldsymbol{v}$.**

4.3.1 Example Problem.

Eigenvalues of Matrix A

Consider a matrix A:

$$A = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$$

The characteristic equation is:

$$\det(A - \lambda I) = \det \begin{pmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{pmatrix} = 0$$

Expanding the determinant:

$$(4-\lambda)(3-\lambda) - 2 \times 1 = 0$$

Simplifying:

$$\lambda^2 - 7\lambda + 10 = 0$$

Solving this quadratic equation:

$$\lambda_1 = 5, \quad \lambda_2 = 2$$

4.3.1 Example Problem.

Eigenvectors of Matrix A

Next, we find the eigenvectors:
For $\lambda_1 = 5$, solve $(A - 5I)v = 0$:

$$(A - 5I) = \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix}$$

Solving the system:

$$\begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0$$

This gives the eigenvector:

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

For $\lambda_2 = 2$, solve $(A - 2I)v = 0$:

$$(A - 2I) = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$$

Solving the system:

$$\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0$$

This gives the eigenvector:

$$v_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Conclusion: For the matrix A, the eigenvalues are $\lambda_1 = 5$ and $\lambda_2 = 2$, with corresponding eigenvectors $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

4.4 Eigenvalue Decomposition.

- **Eigenvalue Decomposition** is a process where a **square matrix is factorized** into
  - its eigenvalues and eigenvectors.
  - Specifically, for a matrix A, if it can be decomposed into a product of three matrices:
    - $\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}$
    - where:
      - $\mathbf{A}$ is the original matrix.
      - $\mathbf{V}$ is the matrix whose columns are the eigenvectors of A.
      - $\boldsymbol{\Lambda}$ is a diagonal matrix whose diagonal entries are the eigenvalues of A.
      - $\mathbf{V}^{-1}$ is the inverse of the matrix V.
- One of the application of Eigenvalue decomposition is **Principal Component Analysis** used for dimensionality reduction purposes.

Thank You.