Math 405 : Learning From Data

Week 1 : Linear Algebra Refresher

Dr. Maher Boudabra

August 26, 2025

Lecture 1: The Language of Linear Algebra

A reality ...

A reality ...

What is Linear Algebra?

Algebra: A formal system of objects and rules for their manipulation.

- **Linear Algebra**: The study of vectors and the rules for adding them and multiplying them by scalars.
- It provides the language to describe **lines, planes, and higher-dimensional analogues** and how they interact.

Central Tenet: *Linearity* - The whole is the sum of its parts. If you know what a transformation does to a set of basic building blocks, you know what it does to everything.

What is a Vector? More Than Just Arrows

Formal Definition: A vector is any object in a set where **addition** and **scalar multiplication** are defined and behave "nicely".

Concrete Examples

Geometric Vectors Directed segments. $\vec{x} + \vec{y} = \vec{z}$, $\lambda\vec{x}$ scales it.

Polynomials e.g., $p(x) = a_0 + a_1 x + a_2 x^2$. $(p+q)(x)$ is a polynomial. $(\lambda p)(x)$ is a polynomial.

Audio Signals A discretized function $s(t)$. The sum of two signals is a signal. Amplifying a signal is scalar multiplication.

Data Points A house's features: $\begin{bmatrix} \text{price} \\ \text{sq. ft.} \\ \#\text{bedrooms} \end{bmatrix}$. We can meaningfully add and scale these.

Our Main Focus: $\mathbb{R}^n$

While the theory is general, our computational workhorse is the space of n-**tuples of real numbers**.

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \middle| \, x_1, x_2, \ldots, x_n \in \mathbb{R} \right\}$$

Operations are defined component-wise:

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}, \quad \lambda \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{bmatrix}$$

This corresponds perfectly to **arrays** in programming languages (NumPy, MATLAB, etc.).

The Central Problem: Systems of Linear Equations

A huge number of problems (across science, engineering, economics, ML) can be formulated as: **Find** $x_1, x_2, \ldots, x_n$ **such that:**

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Real-World Example: Production Planning A company makes products $N_1, \ldots, N_n$ using resources $R_1, \ldots, R_m$.

- To produce one unit of product N_j, it needs a_{ij} units of resource R_i.
- We have b_i total units of resource R_i available.
- **Question:** How many units x_j of each product N_j should we make to use all resources exactly?

This leads directly to the system $\boldsymbol{Ax} = \boldsymbol{b}$.

Matrix Notation: A Revolutionary Compactness

The system of equations is compactly written as:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{\boldsymbol{b}} \qquad \iff \qquad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

Crucial Interpretation: The product $\boldsymbol{A}\boldsymbol{x}$ is a linear combination of the columns of $\boldsymbol{A}$.

$$\boldsymbol{A}\boldsymbol{x} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Solving $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ means finding the coefficients $(x_1, \ldots, x_n)$ that combine the columns of $\boldsymbol{A}$ to form the vector $\boldsymbol{b}$.

Geometric Interpretation of Solutions

In $\mathbb{R}^2$: Each equation $a_{i1}x_1 + a_{i2}x_2 = b_i$ defines a line.

- **No solution**: Lines are parallel.
- **Unique solution**: Lines intersect at one point.
- **Infinitely many**: Lines are identical.

In $\mathbb{R}^3$: Equations define planes. The solution set can be a point, a line, a plane, or empty.

Matrices: The Fundamental Object

Definition: An $(m \times n)$ matrix $\boldsymbol{A}$ is a rectangular array of numbers.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

- a_{ij} is the element in the i-th **row** and j-th **column**.
- A $(1, n)$-matrix is a **row vector**.
- An $(m, 1)$-matrix is a **column vector**.
- The set of all real (m, n)-matrices is denoted $\mathbb{R}^{m \times n}$.

Matrix Operations: Addition and Scalar Multiplication

Addition (Element-wise, matrices must be same size):

$$\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Scalar Multiplication (Element-wise):

$$\lambda \boldsymbol{A} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \cdots \\ \lambda a_{21} & \lambda a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

These operations inherit nice properties from real numbers: associativity, commutativity, distributivity.

Matrix Multiplication: The Non-Obvious Operation

For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{B} \in \mathbb{R}^{n \times p}$, the product $\boldsymbol{C} = \boldsymbol{AB} \in \mathbb{R}^{m \times p}$ is defined by:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

In words: The element c_{ij} is the dot product of the i-th **row** of $\boldsymbol{A}$ and the j-th **column** of $\boldsymbol{B}$.

Crucial: $\boldsymbol{AB}$ is only defined if the number of **columns** of $\boldsymbol{A}$ equals the number of **rows** of $\boldsymbol{B}$.

Properties of Matrix Multiplication

- **Associative**: $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$
- **Distributive**: $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$
- **NOT Commutative**: $\boldsymbol{AB} \neq \boldsymbol{BA}$ in general! Order matters immensely.
- **Identity Matrix**: $\boldsymbol{I}_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$. For any $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A}$.

Why is it defined this way?

Matrix multiplication is defined so that it represents the **composition of linear maps**. If $\boldsymbol{A}$ represents transformation T and $\boldsymbol{B}$ represents transformation S, then $\boldsymbol{AB}$ represents $T(S(\boldsymbol{x}))$.

Special Matrices: Transpose and Inverse

Transpose: Flip rows and columns. $(\boldsymbol{A}^\top)_{ij} = a_{ji}$.

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

Properties: $(\boldsymbol{A}^\top)^\top = \boldsymbol{A}$, $(\boldsymbol{AB})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top$. **Inverse**: For a **square** matrix $\boldsymbol{A}$, if there exists

$\boldsymbol{B}$ such that $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$, then $\boldsymbol{B}$ is the inverse $\boldsymbol{A}^{-1}$.

- Not all matrices are invertible. Those that are not are called **singular**.
- If it exists, it is unique, and $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, $(\boldsymbol{A}^\top)^{-1} = (\boldsymbol{A}^{-1})^\top$.

Symmetric Matrix: $\boldsymbol{A} = \boldsymbol{A}^\top$. Very common in applications (e.g., covariance matrices).

Summary - Lecture 1

- Linear Algebra studies **vectors** (objects that can be added/scaled) and **linear maps**.
- The central problem is solving **systems of linear equations** $\boldsymbol{Ax} = \boldsymbol{b}$.
- **Matrices** provide a compact notation and represent both data (collections of vectors) and linear transformations.
- Matrix multiplication is a powerful but non-commutative operation.
- The **inverse** $\boldsymbol{A}^{-1}$ (if it exists) solves the system: $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

Next: How do we **systematically solve** these systems, especially when $\boldsymbol{A}$ is not invertible or not even square?

Lecture 2: Solving Systems and Vector Spaces

The Structure of All Solutions

The complete solution to $\boldsymbol{Ax} = \boldsymbol{b}$ has a beautiful structure:

$$\text{General Solution} = \text{Particular Solution} + \text{Homogeneous Solution}$$

Algorithm to Solve $Ax = b$

1. Find **any one** specific solution $\boldsymbol{x}_p$ to $\boldsymbol{Ax} = \boldsymbol{b}$.
2. Find **all** solutions $\boldsymbol{x}_h$ to the **homogeneous** system $\boldsymbol{Ax} = 0$.
3. The general solution is $\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_h$.

Why does this work? Because $\boldsymbol{A}(\boldsymbol{x}_p + \boldsymbol{x}_h) = \boldsymbol{Ax}_p + \boldsymbol{Ax}_h = \boldsymbol{b} + 0 = \boldsymbol{b}$.

Gaussian Elimination: The Great Algorithm

How do we find these solutions? Use **elementary row operations** on the **augmented matrix** $[\boldsymbol{A}|\boldsymbol{b}]$:

1. **Swap** two rows.
2. **Multiply** a row by a non-zero scalar.
3. **Add** a multiple of one row to another row.

These operations preserve the solution set! Our goal is to transform the matrix into a simpler form.

Row Echelon Form (REF) and Rank

A matrix is in **REF** if:

- All rows of all zeros are at the bottom.
- The first non-zero number (the **pivot**) of a row is always strictly to the right of the pivot above it.

Example

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

Pivots are in **bold**. The number of pivots is the **rank** of the matrix, $\mathrm{rank}(\boldsymbol{A})$.

Rank Theorem: For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\mathrm{rank}(\boldsymbol{A}) \leq \min(m, n)$. It is the **dimension of the column space** (number of independent columns).

Using REF to Find Solutions

Once in REF, we can solve by **back-substitution**.

Example

Solve from the augmented REF matrix:

$$\left[\begin{array}{ccc|c} \mathbf{1} & 2 & -1 & 3 \\ 0 & \mathbf{1} & 3 & 1 \\ 0 & 0 & \mathbf{2} & 4 \end{array}\right]$$

1. Row 3: $2x_3 = 4 \Rightarrow x_3 = 2$
2. Row 2: $x_2 + 3(2) = 1 \Rightarrow x_2 = -5$
3. Row 1: $x_1 + 2(-5) - (2) = 3 \Rightarrow x_1 = 15$

Unique solution: $(15, -5, 2)$.

The Minus-1 Trick for $\ker(\boldsymbol{A})$

A mechanical way to find a basis for the null space from the **reduced REF** (RREF).

1. Bring $\boldsymbol{A}$ to RREF.
2. For each column **without** a pivot, add a row $[0 \dots 0 - 1 0 \dots 0]$ where the -1 is in the position of that free variable column.
3. The columns containing the -1s form a **basis** for the null space.

Example

For $\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix}$ (RREF), columns 2 and 5 are non-pivot.

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \Rightarrow \ker(\boldsymbol{A}) = \operatorname{span}\left\{ \begin{bmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix} \right\}$$

Calculating the Inverse via Elimination

To find $\boldsymbol{A}^{-1}$, solve $\boldsymbol{AX} = \boldsymbol{I}$.

1. Form the augmented matrix $[\boldsymbol{A}|\boldsymbol{I}_n]$.
2. Perform Gaussian Elimination until the left block is in **Reduced REF** $(\boldsymbol{I}_n)$.
3. If successful, the right block is $\boldsymbol{A}^{-1}$: $[\boldsymbol{I}_n|\boldsymbol{A}^{-1}]$.

Example

Find inverse of $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \sim \cdots \sim \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 1.5 & -0.5 \end{array}\right] \Rightarrow \boldsymbol{A}^{-1} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix}$$

Abstracting the Concept: Vector Spaces

We now define the general setting where all this lives. **Definition:** A **real vector space** V is a set of objects (vectors) with two operations:

- **Vector addition** $+ : V \times V \to V$
- **Scalar multiplication** $\cdot : \mathbb{R} \times V \to V$

that satisfy 8 axioms (associativity, commutativity, distributivity, existence of $\mathbf{0}$ and $-\boldsymbol{x}$, etc.).

Key Examples

- $\mathbb{R}^n$, $\mathbb{R}^{m \times n}$
- Polynomials of degree $\leq n$: $\{a_0 + a_1 x + \ldots + a_n x^n\}$
- Continuous functions on $[a, b]$: $C([a, b])$
- The solution set of $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ (The **null space**)

Subspaces

Definition: A subset $U \subseteq V$ of a vector space is a **subspace** if it is itself a vector space under the same operations.

Theorem (Subspace Test)

$U \subseteq V$ is a subspace if and only if:

1. *$U \neq \emptyset$ (usually check $\mathbf{0} \in U$).*
2. ***Closed under addition**: $\forall \boldsymbol{u}, \boldsymbol{v} \in U, \boldsymbol{u} + \boldsymbol{v} \in U$.*
3. ***Closed under scalar multiplication**: $\forall \lambda \in \mathbb{R}, \boldsymbol{u} \in U, \lambda \boldsymbol{u} \in U$.*

Example

- The set of all vectors in $\mathbb{R}^3$ whose components sum to zero: $\{(x, y, z) \mid x + y + z = 0\}$.
- The span of any set of vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$.
- **Crucially:** The solution set of $\boldsymbol{A}\boldsymbol{x} = 0$ is a subspace. The solution set of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ ($\boldsymbol{b} \neq 0$) is **not** a subspace (it lacks the zero vector).

Summary - Lecture 2

- The solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has the form $\boldsymbol{x}_p + \boldsymbol{x}_h$.
- **Gaussian Elimination** is the fundamental algorithm for solving systems, finding rank, and calculating inverses.
- The **rank** of a matrix is the number of independent rows/columns.
- A **Vector Space** is an abstract set where addition and scaling make sense.
- A **Subspace** is a subset that is closed under these operations.

Next: How to describe vector spaces efficiently? How do we measure their "size"? How do we define transformations between them?

Lecture 3: Basis, Linear Maps, and Applications

Linear Combinations and (In)dependence

Linear Combination: A sum of scaled vectors: $\boldsymbol{v} = \lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_k \boldsymbol{v}_k$.
The set of all linear combinations is the **span**: $\mathrm{span}(\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\})$. **Linear Dependence**: A

set of vectors is **linearly dependent** if one vector can be written as a linear combination of the others. Equivalently, there exist scalars (not all zero) such that:

$$\lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_k \boldsymbol{v}_k = \mathbf{0}$$

Linear Independence: The set is **linearly independent** if the *only* solution to the above equation is $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$.

Example

$\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \end{bmatrix}$ are dependent ($2^{\text{nd}} = 3 \times 1^{\text{st}}$). $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are independent.

Basis and Dimension

Basis: A **basis** $\mathcal{B}$ for a vector space V is a set of vectors that is:

1. **Linearly independent**.
2. A **generating set** for V (i.e., $\operatorname{span}(\mathcal{B}) = V$).

Dimension: The number of vectors in any basis for V is called the **dimension** of V, denoted $\dim(V)$.

Example

The **standard basis** for $\mathbb{R}^3$ is:

$$\mathcal{E} = \left\{ \boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$\dim(\mathbb{R}^3) = 3$. But this is not the only basis!

Coordinates

Let $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be a basis for V.
For any vector $\boldsymbol{v} \in V$, there exists a **unique** set of scalars $(\alpha_1, \ldots, \alpha_n)$ such that:

$$\boldsymbol{v} = \alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2 + \cdots + \alpha_n \boldsymbol{b}_n$$

The vector $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{R}^n$ is called the **coordinate vector** of $\boldsymbol{v}$ with respect to the basis $\mathcal{B}$.

Linear Mappings

Definition: A function $\Phi : V \to W$ between vector spaces is a **linear mapping** (or **linear transformation**) if it preserves the vector space structure:

1. $\Phi(\boldsymbol{u} + \boldsymbol{v}) = \Phi(\boldsymbol{u}) + \Phi(\boldsymbol{v})$ (Preserves addition)
2. $\Phi(\lambda \boldsymbol{v}) = \lambda \Phi(\boldsymbol{v})$ (Preserves scalar multiplication)

This can be combined into one condition: $\Phi(\lambda \boldsymbol{u} + \psi \boldsymbol{v}) = \lambda \Phi(\boldsymbol{u}) + \psi \Phi(\boldsymbol{v})$.

Example

- **Differentiation**: $\frac{d}{dx}$: Polynomials $\to$ Polynomials is linear.
- **Matrix Multiplication**: For a fixed matrix $\boldsymbol{A}$, the map $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ defined by $\Phi(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ is linear.
- **Rotation** in the plane is a linear transformation.

The Matrix of a Linear Map

This is the crucial connection between abstract linear maps and concrete matrix computation.
Let $\Phi : V \to W$ be linear.
Let $\mathcal{B}_V = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be a basis for V.
Let $\mathcal{B}_W = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m\}$ be a basis for W.
For each basis vector $\boldsymbol{b}_j$ in V, express its image under Φ in the basis of W:

$$\Phi(\boldsymbol{b}_j) = a_{1j}\boldsymbol{c}_1 + a_{2j}\boldsymbol{c}_2 + \cdots + a_{mj}\boldsymbol{c}_m$$

The **transformation matrix** $\boldsymbol{A}_\Phi$ is the $m \times n$ matrix whose j-th **column** is the coordinate vector of $\Phi(\boldsymbol{b}_j)$ relative to $\mathcal{B}_W$.
Then, for any $\boldsymbol{x} \in V$, if $\hat{\boldsymbol{x}}$ is its coordinate vector (w.r.t $\mathcal{B}_V$), the coordinate vector of $\Phi(\boldsymbol{x})$ (w.r.t $\mathcal{B}_W$) is:

$$\widehat{\Phi(\boldsymbol{x})} = \boldsymbol{A}_\Phi \hat{\boldsymbol{x}}$$

Image and Kernel (Range and Null Space)

For a linear map $\Phi : V \to W$:

- **Image/Range**: $\mathrm{Im}(\Phi) = \{\Phi(\boldsymbol{v}) \in W \mid \boldsymbol{v} \in V\}$. *What can Φ produce?*
- **Kernel/Null Space**: $\ker(\Phi) = \{\boldsymbol{v} \in V \mid \Phi(\boldsymbol{v}) = \mathbf{0}_W\}$. *What does Φ send to zero?*

Both $\mathrm{Im}(\Phi) \subseteq W$ and $\ker(\Phi) \subseteq V$ are **subspaces**.

The Rank-Nullity Theorem

Theorem (Fundamental Theorem of Linear Maps)

Let $\Phi : V \to W$ be a linear map, and assume V is finite-dimensional. Then:

$$\dim(\ker(\Phi)) + \dim(\mathrm{Im}(\Phi)) = \dim(V)$$

Terminology:

- $\dim(\mathrm{Im}(\Phi))$ is the **rank** of Φ.
- $\dim(\ker(\Phi))$ is the **nullity** of Φ.

Implications:

- If $\dim(V) > \dim(W)$, Φ cannot be injective (it must "crush" $\dim(V) - \dim(W)$ dimensions down to zero).
- If $\dim(V) = \dim(W)$, then injectivity $\iff$ surjectivity $\iff$ bijectivity.

Affine Spaces: "Vector Spaces That Don't Pass Through Origin"

The solution set to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ (with $\boldsymbol{b} \neq 0$) is not a subspace. It is an **affine space**. **Definition:** A subset L of a vector space V is an **affine subspace** if it can be written as

$$L = \boldsymbol{x}_0 + U = \{\boldsymbol{x}_0 + \boldsymbol{u} \mid \boldsymbol{u} \in U\}$$

where $\boldsymbol{x}_0 \in V$ is a fixed **support point** and $U \subseteq V$ is a **subspace** (the **direction** or **parallel subspace**).

Example

- A **line** in $\mathbb{R}^3$ not through the origin: $L = \boldsymbol{p} + \mathrm{span}\{\boldsymbol{d}\}$.
- A **plane** in $\mathbb{R}^3$ not through the origin: $L = \boldsymbol{p} + \mathrm{span}\{\boldsymbol{u}, \boldsymbol{v}\}$.
- The solution set of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is $L = \boldsymbol{x}_p + \ker(\boldsymbol{A})$.

Summary - Lecture 3 & Course

- A **Basis** provides a coordinate system for a vector space. Its size is the **Dimension**.
- **Linear Maps** are the structure-preserving functions between vector spaces.
- Every linear map can be represented by a **Matrix**, once bases are chosen.
- The **Image** (what is hit) and **Kernel** (what is crushed) are fundamental.
- The **Rank-Nullity Theorem** $\dim(V) = \dim(\ker) + \dim(\mathrm{Im})$ is a cornerstone.
- Solutions to inhomogeneous systems form **Affine Spaces**.

Where to next? This is the *algebraic* foundation. The *geometric* foundation (lengths, angles, distances, orthogonality, projections) is built on top of this using **inner products**, which is the starting point for many machine learning algorithms (e.g., linear regression, PCA).