Computer Vision

Dr. Syed Faisal Bukhari

Associate Professor

Department of Data Science

Faculty of Computing and Information Technology

University of the Punjab

Reference

These notes are based on

- Dr. Matthew N. Dailey's course: AT70.20: Machine Vision for Robotics and HCI

Following are the fundamental problems in computer vision:

1. 2D homography: Given a set of points x_i in $\mathbb{P}^2$ and corresponding points x_i' in $\mathbb{P}^2$ find a homography taking each x_i to x_i'

2. 3D to 2D camera projection: Given a set of points X_i in 3D space and corresponding points x_i in an image, find the 3D to 2D projective mapping taking each X_i to x_i .

3. Fundamental matrix computation: Given a set of points x_i in one image and a set of corresponding points x_i' in another image, find the **fundamental matrix F** relating the two images.

4. Trifocal Tensor computation: Given a set of **point correspondences** $x_i \leftrightarrow x_i' \leftrightarrow x_i''$ across **three images**, compute the trifocal tensor T_i^{jk} relating points or lines in three views.

Projective Geometry and Transformations of 3D

- The **projective space** $\mathbb{P}^3$ augments **Euclidean 3-space** by adding a **set of ideal points**, all lying on the **plane at infinity**, denoted as $\boldsymbol{\pi}_\infty$

- This is the analogous to **line at infinity** l_∞ in $\mathbb{P}^2$.

- **Parallel lines**, and now **parallel planes**, **intersect** on $\boldsymbol{\pi}_\infty$.

- As expected, **homogeneous coordinates** once again play an important role, with all **dimensions increased by one**.

- **For example**, any **two parallel lines** always **meet on the 2D projective plane**, even though they may **not intersect in 3D.**

Points and projective transformations

○ Representation in 3-Space

- **A** point **X or $\vec{X}$** in **3-space** is represented in **homogeneous coordinates** as a **4-vector**.

○ Homogeneous Vector Definition

- Specifically, the **homogeneous vector is:**

  $$\vec{X}^T = (X_1, \quad X_2, \quad X_3, \quad X_4)^T$$

- **When** $X_4 \neq 0$, it represents the Euclidean point $\mathbb{R}^{\mathbf{3}}$:

  $$(X, \quad Y, \quad Z)^{\boldsymbol{T}}$$

  where X = $\frac{X_1}{X_4}$, Y = $\frac{X_2}{X_4}$, Z = $\frac{X_3}{X_4}$

These are called **inhomogeneous coordinates** in $\mathbb{R}^{\mathbf{3}}$.

Points and projective transformations

○Homogeneous Representation Example

- For example, a homogeneous representation of a Euclidean point $\vec{X}^T = (X, \quad Y, \quad Z)^{\boldsymbol{T}}$
 is $\vec{X}^T = (X, \quad Y, \quad Z, \quad 1)^T$.

○Points at Infinity

- **Homogeneous points** with $X_4 = 0$ represent **points at infinity.**
- **These points** lie on the **plane at infinity π_∞** in projective space.

Points and Projective Transformations in $\mathbb{P}^3$

- A **projective transformation** acting on $\mathbb{P}^3$ is a linear transformation on **homogeneous 4-vectors** represented by a **non-singular 4 × 4 matrix**:

$$X'_{4\times1} = H_{4\times4}X_{4\times1}$$

- The **matrix H**, which defines the transformation, is **homogeneous** and has **15 degrees of freedom**.

- The degrees of freedom of **matrix H** follows from the **16 elements** of the matrix **less one for overall scaling**.

3D Projective Geometry

- We will explore **transformations** involving **3D spaces, 3D objects, or 3D sets of points.**

- Imagine that on the **left side** we have one **3D space**, and on the **right side another 3D space**.

- There are **transformations** connecting **these two spaces**. To understand them, we must visualize these **3D spaces** as **embedded** within a **higher-dimensional space**. Essentially, we need to think in **4D** to **comprehend transformations between 3D spaces**.

3D Projective Geometry

- We will generalize notion of **homogeneous points** from 2D to 3D. A **3D point** is represented by a **homogenous 4 vector.** Given a homogeneous vector

$$\vec{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

In order to obtain a point in $\mathbb{R}^3$, we need to **normalize**

$$X = \frac{x_1}{x_4}, \quad Y = \frac{x_2}{x_4}, \quad \text{and} \quad Z = \frac{x_3}{x_4}$$

3D Projective Geometry

- We see things are a little different in **3D** than **2D**.

- In 2D, we found, if points are in **homogenous representations**, then **points** and **lines** are **dual** to each other.

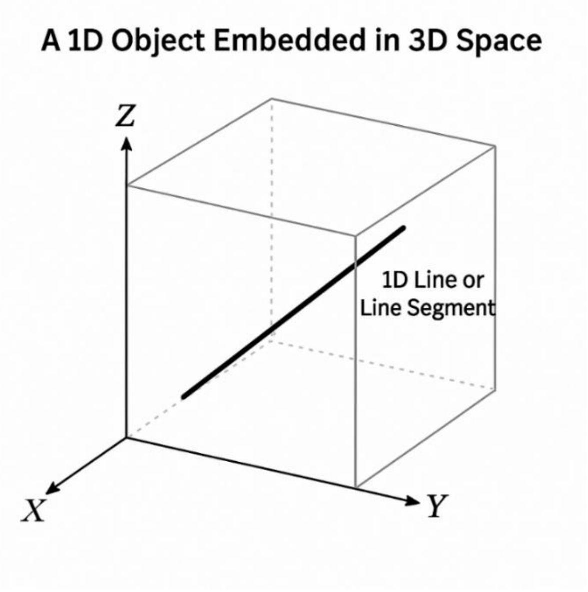
- In 3D, we will find that **points** and **lines** don't have a special relationship with each other.

- In fact, **lines in 3D** are **very difficult to deal.**

What is a line in 3D?

- It refers to either a line segment or an infinite line, conceptualized as a **one-dimensional (1D)** entity embedded within **three-dimensional (3D) space**

- Think of it as a **1D object** — like a **line** or a **line segment** — that exists inside a **3D world**.



Describing a 3D Line

- A 3D line can be defined by two points (take two points on a 3D line), **giving 6 elements**.
- Despite having 6 elements, a 3D line actually **has only 4 degrees of freedom**. Finding the **minimal representation** is essential.
- **Without a minimal form**, it becomes **challenging to estimate lines in 3D**, especially in statistical and algebraic estimation.
- Just like the **duality** between **points and lines** in 2D, there exists a similar **duality** between **points and planes in 3D**.

3D Homographies

- In **2D**, a **line** is considered a **1D object** embedded in **2D space**. Similarly, in **3D space**, a **plane** is a **2D object** embedded in **3D space**. This loss of **one dimension** introduces the **concept of duality.**

- Previously, we discussed **2D homographies**. **3D homographies** extend these ideas. When working with projective transformations, the terms **homographies** and **projective transformations** are sometimes used interchangeably, although they technically refer to same concept.
- The relationship between the coordinates is given by:

  X' = H X or $X'_{4\times 1} = H_{4\times 4}X_{4\times 1}$

3D Homographies

- In lecture notes, we will use $\vec{x}$ to denote **2D points** and $\vec{X}$ to denote 3D points.

- We have a corresponding 3D point $\vec{X}\,'$. We have used **"="** sign in **homogeneous sense**. It means proportional to. The result is a 3D point $\vec{X}\,'$ by some **homography**. This homography matrix or projective transformation is how big:

$$\vec{X}\,'_{4\times1} = H_{4\times4}\,\vec{X}_{4\times1}$$

Planes in 3D Space

- A plane in 3D space is represented by the equation:

  $\pi_1 X + \pi_2 Y + \pi_3 Z + \pi_4 = 0$

- Here, $\pi_1, \pi_2, \pi_3,$ and π_4 are constants (coefficients).

  Any **point (X, Y, Z)** that satisfies **this equation** lies on **the plane**. This is **similar** to the **equation of a line** in 2D.

- **In 2D, a line is given by:**

  $ax + by + c = 0$

- Notice the similarity between the two forms:

  $ax + by + c = 0$

  $\pi_1 X + \pi_2 Y + \pi_3 Z + \pi_4 = 0$

- This resemblance helps us intuitively understand the connection between **planes in 3D** and **lines in 2D**.

Visual Comparison: 2D Line vs 3D Plane

Line in 2D

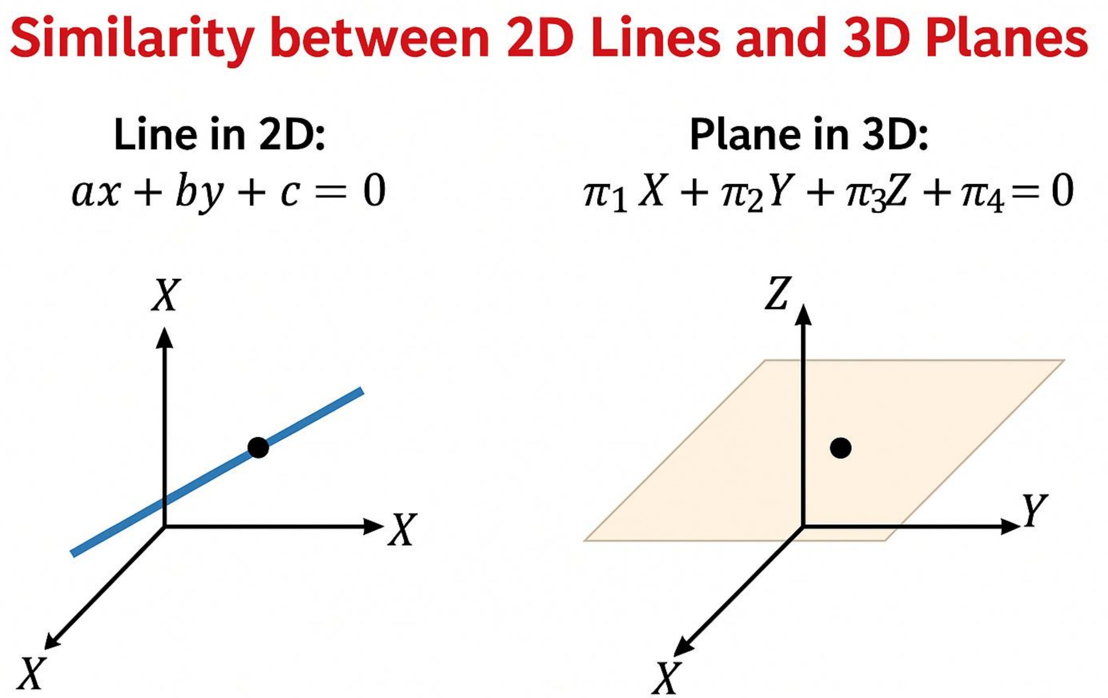
Equation: $ax + by + c = 0$

Graph: A straight line on a 2D plane

Plane in 3D

Equation: $\pi_1 X + \pi_2 Y + \pi_3 Z + \pi_4 = 0$

Graph: A flat surface in 3D space



Planes in 3 space

A plane in 3-space may be written as

$\pi_1 X + \pi_2 Y + \pi_3 Z + \pi_4 = 0$ ----------------(1)

- Clearly this equation is **unaffected** by multiplication by a **non-zero scalar**. It follows that a **plane has 3 degrees of freedom in 3-space**

- The **homogeneous representation of the plane** is the **4-vector**

$$\vec{\pi} = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{bmatrix} \text{ or } (\pi_1, \quad \pi_2, \quad \pi_3, \quad \pi_4)^T$$

Planes in 3 space

Homogenizing (1) this by the replacements:

$$X \longmapsto \frac{X_1}{X_4}, Y \longmapsto \frac{X_2}{X_4}, Z \longmapsto \frac{X_3}{X_4}$$

$$\pi_1 \frac{X_1}{X_4} + \pi_2 \frac{X_2}{X_4} + \pi_3 \frac{X_3}{X_4} + \pi_4 = 0$$

$$\Rightarrow \boldsymbol{\pi_1 X_1 + \pi_2 X_2 + \pi_3 X_3 + \pi_4 X_4 = 0}$$

or more concisely

$$\boldsymbol{\pi^T \mathrm{X} = 0}$$

which expresses that the point **X** is on the plane $\boldsymbol{\pi}$.

Planes in 3 space

- **Recall:** We know if a point lies on a line. We take the **dot product** of the **coefficient vector** with the **homogeneous representation** of a point, and we just add the value 1 to the third component. If the **dot product is equal to zero**, then we say that the **point is on the line**.

- The form of the equation $\pi_1 X + \pi_2 Y + \pi_3 Z + \pi_4 = 0$ is **same** as the **equation of line** ax + by + c = 0

- We have a **homogeneous representation** of a **point in 3 space**. We take the **dot product** with the **coefficient vector describing the plane** i.e.,

  $\vec{\pi} = (\pi_1, \quad \pi_2, \quad \pi_3, \quad \pi_4)^T$

Planes in 3 space

- It means, if we get the **dot product equal** to **zero** then it means our **3D point lies on the plane** $\vec{\pi}$.

- In other words, if $\vec{X}$ is on $\vec{\pi}$ = $(\pi_1, \quad \pi_2, \quad \pi_3, \quad \pi_4)^T$ then $\vec{X}^{\,T}\,\vec{\pi} = 0$.

- That's the way we tell if a particular **point** $\vec{X}$ lies on the plane $\vec{\pi}$.

How can we determine what is the plane that contains the 3D points?

- **Recall:** In 2D, if want to know what is the line that contains the two 2D points. We take the **cross product of these 2D** points in order to get the **equation of line** that contains these two 2D points.
- In **case of 3D**, we need **three 3D points** to determine a plane. We take a **4D homogenous vector** representing a **3D point**. We figure out what is the plane that spans these points
- If $\vec{X}$ is on $\vec{\pi}$ then $\vec{X}^{T}\vec{\pi} = 0$
- Similarly, if **three 3D points** $\vec{X}_1, \vec{X}_2, \vec{X}_3$ are on π then

$$\vec{X}_1^T\vec{\pi} = 0$$

$$\vec{X}_2^T\vec{\pi} = 0$$

$$\vec{X}_3^T\vec{\pi} = 0$$

How can we determine what is the plane that contains the 3D points?

In compact form, we can write as

$$\begin{bmatrix} X_1^T \\ X_2^T \\ X_3^T \end{bmatrix}_{3\times4} \vec{\pi}_{4\times1} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}_{3\times1}$$

OR

$$\underbrace{\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 \\ X_2 & Y_2 & Z_2 & 1 \\ X_3 & Y_3 & Z_3 & 1 \end{bmatrix}_{3\times4}}_{\text{Three 3D points}} \underbrace{\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{bmatrix}_{4\times1}}_{\text{Parameters of plane}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}_{3\times1}$$

These **three 3D points** are supposed to lie on the **plane** π. If we transpose the vector, now we get a 4-vector dot product with a 4-vector describing the plane π should be equal to zero. We get

$\vec{X}_1^T \vec{\pi} = 0$

$\vec{X}_2^T \vec{\pi} = 0$

$\vec{X}_3^T \vec{\pi} = 0$

We saw exactly the same problem when we are trying to **estimate** a **homography** from a set of **4 correspondences between 2 views**.

- **Recall:** A 2D homography is **9-dimensional object**. It has 9 elements. But it is a homogenous object. So, it has **8 degrees of freedom**. We have seen, we can build a linear system using **4 points correspondences**. In this system, we have 8 equations and 9 unknowns. This system provides 8 equations for the 8 degrees of freedom in a homography matrix (since one parameter is fixed due to scale)

- Similarly, if we stack up 3 points lying on a plane, we obtain a 3×4 matrix, leading to 3 equations in 4 unknowns — representing the homogeneous coordinates of the plane $\pi \in \mathbb{P}^3$

$$\begin{bmatrix} X_1^T \\ X_2^T \\ X_3^T \end{bmatrix}_{3\times 4}$$

Review

Full rank:

- If our matrix is an **$m \times n$** matrix with **$m < n$**, then it has **full rank** when its **m rows are linearly independent**.

- If **$m > n$**, the matrix has **full rank** when its **n columns are linearly independent**.

- If **$m = n$**, the matrix has **full rank** either when its **rows or its columns are linearly independent** (when the rows are linearly independent, so are its columns in this case).

Case 1: m < n

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

A is a $\mathbf{2 \times 3}$ matrix. The **2 rows** are **linearly independent**

⇒ Full row rank.

Case 2: m > n

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

A is a **3 x 2 matrix**. The **2 columns are linearly independent**

⇒ Full column rank.

Case 3: m = n (Full Rank)

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

A is a 2 x 2 matrix with **full rank** since its rows (and thus columns) are **linearly independent**.

Case 4: m = n (Not Full Rank)

A = $\begin{bmatrix} \mathbf{1} & \mathbf{2} \\ \mathbf{2} & \mathbf{4} \end{bmatrix}_{2\times 2}$

A is 2 x 2 matrix has linearly dependent rows

$\Rightarrow$ Rank < 2

$\Rightarrow$ Not full rank.

Is π Homogeneous?

- Yes, π is homogeneous.

- The equation is: $\pi_1 X + \pi_2 Y + \pi_3 Z + \pi_4 = 0$

- If we multiply the entire equation by **any nonzero scalar**, we are **essentially scaling the coefficients** $(\pi_{1,}\ \pi_2,\ \pi_{3,}\ \pi_4)$ by that scalar.

- Since the **original equation equals zero**, scaling it by any factor still results in zero. **Thus, any multiple of this equation remains zero.**

- Any vector π that satisfies this system of **linear equations** $(X_1^T, \quad X_2^T, \quad X_3^T)_{3\times 4}^T$ is a representation of the **plane** that contains these **three points**.

- **How we solve the problem.** How we solve it last time? The **null vector of this matrix**. Similarly take the **null space** of $(X_1^T, \quad X_2^T, \quad X_3^T)_{3\times 4}^T$ matrix.

- The **null space** is when you get a null vector or set of null vectors for a particular matrix that gives you vectors that **spans** the null space of that matrix. That means the set of all vectors **π** when multiply with the matrix $(X_1^T, \quad X_2^T, \quad X_3^T)_{3\times 4}^T$ on left with vector π on your right

Understanding the Plane Through Three Points

- **Any vector** π that satisfies the system of linear equations $(X_1^T, \quad X_2^T, \quad X_3^T)_{3\times4}^T$ represents the **plane passing through these three points.**
- How do we solve this?

  We compute the null vector of the matrix. Specifically, we take the null space of the matrix $(X_1^T, \quad X_2^T, \quad X_3^T)_{3\times4}^T$

- The null space consists of all vectors π such that $(\boldsymbol{X_1^T}, \quad \boldsymbol{X_2^T}, \quad \boldsymbol{X_3^T})_{\boldsymbol{3\times4}}^{\boldsymbol{T}} \cdot \pi = 0$
- This gives the directions that span the null space, i.e., the set of solutions that lie in the plane.

Null Space and Unique Solution

- The null space of a matrix contains all vectors that map to the zero vector.
  For a **full-rank matrix**, the null space is trivial—it includes only the zero vector. No non-zero vector satisfies Ax = 0 in this case.

- **So, when do we get a unique solution?**
  - If the matrix formed by stacking $(X_1^T, \quad X_2^T, \quad X_3^T)^T$ vertically has rank 3, its null space is one-dimensional. This guarantees a unique direction for the solution space.
  - Therefore, the rank must be 3 to ensure a unique solution

Rank 3 Condition and Linear Dependence

- **What does it mean for the matrix to have rank 3?**
  - It means the matrix has three linearly independent rows.
  - If the 4D representations of the points are linearly independent, the system is stable.

- **When does linear dependence occur?**
  - When all three points lie on the same line
  - When any two of the points are identical
  - In these cases, the rank drops and the system loses uniqueness.

- Therefore, if the points $(X_1^T, \quad X_2^T, \quad X_3^T)^T$ do not lie on the same line, the matrix has rank 3 and we **get a unique solution**.

Plane from Null Space of Stacked Points

- If the matrix $(X_1^T, \quad X_2^T, \quad X_3^T)^T$ has rank 2, then any solution to $\boldsymbol{X^T \pi}$ = 0 describes a **plane containing the line** that passes through the points $\boldsymbol{X_1, X_{2,} X_3}$.
- But we don't need **three 3D points** to find such a **plane**. If you stack any **two 3D points** into a 2×4 matrix and compute its null space, you will obtain **two basis vectors.**
- Any linear combination of these two basis vectors describes a plane that contains the **line formed by those two points**.

Null Space and Plane Construction in 3D

Question: How can we get those planes?

Ans: If $\vec{X}_1$ and $\vec{X}_2$ are on a **3D line L**, then the **null** $\left(\begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix}\right)$ is the set of all vectors $\boldsymbol{\pi}$ **describing a plane that contains L**

Note: If $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ span the null space of $\overrightarrow{\boldsymbol{X}}$, then any **linear combination** $\boldsymbol{\pi}' = \alpha\boldsymbol{\pi}_1 + \beta\boldsymbol{\pi}_2$ will also satisfies $\overrightarrow{\boldsymbol{X}}\boldsymbol{\pi}' = \vec{\mathbf{0}}$

Describing a Plane Using Points and Null Space

- There are **multiple ways** to define a **plane** through a set of points. One approach uses the **null space**, which is especially useful in MATLAB.
- Construct vectors from **three points**: $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$. Use the **cross product** to find the **plane passing through these points**.
- This method uses the inhomogeneous form of a point:

  $\boldsymbol{X} = (X, Y, Z)^T$.
- In MATLAB or Octave, the null space of the constructed matrix gives the plane equation directly.
- This method connects **points and planes** in 3D projective space $\mathbb{P}^3$.

What is a Pencil of Planes?

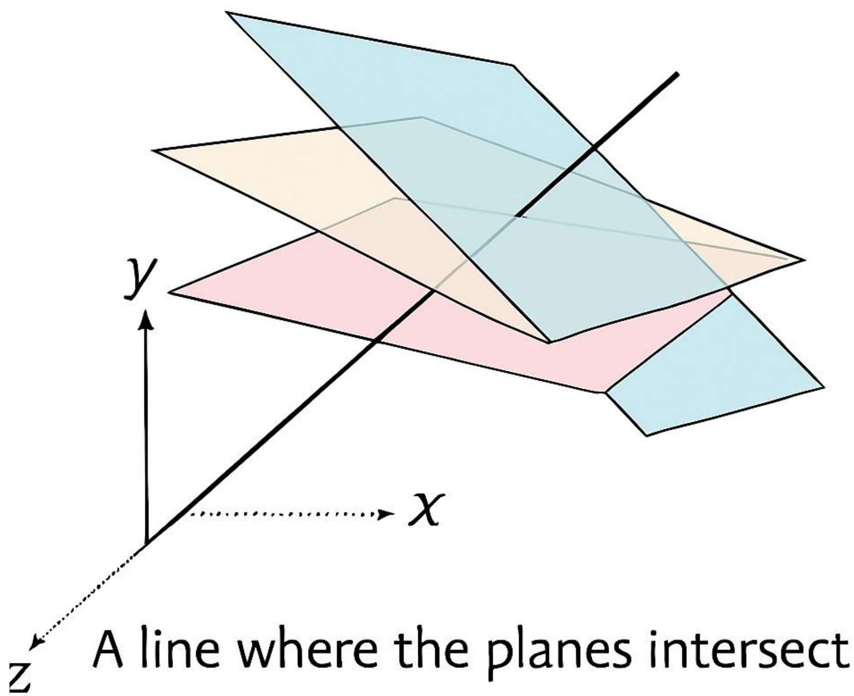
- Imagine a number of points lying along a line in 3D space.

- Now consider all the planes that contain this line:
  - **One plane** passes through **the line**
  - Another **different plane** also passes through **the same line, and so on.**

- When we take the collection of all such planes, we call this a **pencil of planes** through the line.

- In short: All the planes that contain the same line form a **pencil of planes.**

Pencil of planes



Recall: Null Space, Right Null Space, and Left Null Space

- **Null Space (Right Null Space):**
  - Set of all vectors x such that $A \cdot x = 0$
  - Also called the right null space
  - Subspace of $\mathbb{R}^n$ (n = number of columns of A)
  - Represents vectors mapped to zero by A

- **Right Null Space:**
  - Just another name for the null space
  - Equation: $A \cdot x = 0$
  - Subspace of $\mathbb{R}^n$

Recall: Null Space, Right Null Space, and Left Null Space

- **Left Null Space:**
  - Set of all vectors y such that $A^T \cdot y = 0$
  - Subspace of $\mathbb{R}^m$ (m = number of rows of A)
  - Represents vectors orthogonal to rows of A

- **Summary:**
  - Null space and right null space are the same — both refer to x such that $A \cdot x = 0$.
  - Left null space refers to y such that $A^T \cdot y = 0$ (i.e., vectors orthogonal to the rows of A).

Null Space and Left Null Space

A = $\begin{bmatrix} \mathbf{1} & \mathbf{2} \\ \mathbf{3} & \mathbf{4} \end{bmatrix}_{\mathbf{2\times 2}}$

Null Space (Right Null Space):

Find x such that A·x = 0

Let x = $[2 \quad -1]^T \Rightarrow$ A·x = 0 ⇒ x is in Null(A)

Left Null Space:

Find y such that A^T·y = 0

Let y = $[2 \quad -1]^T \Rightarrow$ A^T·y = 0 ⇒ y is in Null(A^T) $[2 \quad -1]^T$

Geometric View:

- Right null space lies in input space (domain)
- Left null space lies in output space (codomain)
- Both are orthogonal complements to the row/column spaces