Vectors & Matrices

CSC4601 Theory of Machine Learning

Today’s Plan

- In this lecture, we’ll review the basics of linear algebra:
  - Vectors & their geometric interpretation
  - Dot products & vector norms
  - Matrices & operations on matrices

- We will also revisit the equation of Line (Plane, Hyperplane) & the definition of normal vectors.

Vectors & Matrices in Machine Learning

- In machine learning, we use vectors and matrices to organize data: see them as data containers/structures.

- They make notation much easier and more concise: Python supports the use of vector notations which can make computations more efficient.

- We can use many tools from linear algebra in order to perform training, prediction, dimensionality reduction.

Vectors

Vectors

- A vector is an array of numerical values. It collects an ordered list of numbers to represent one entity or one mathematical object.
- Vectors are usually represented in one of the following form:

$$\begin{bmatrix} 5 \\ 2 \\ 6 \end{bmatrix}, \begin{pmatrix} 5 \\ 2 \\ 6 \end{pmatrix}, (5, 2, 6)$$

- Notations:
  - $\boldsymbol{v} \in \mathbb{R}^2$: $\boldsymbol{v}$ a vector with 2 real entries
  - $\boldsymbol{v} \in \mathbb{R}^3$: $\boldsymbol{v}$ a vector with 3 real entries
  - $\boldsymbol{v} \in \mathbb{R}^n$: $\boldsymbol{v}$ a vector with n real entries

Vectors

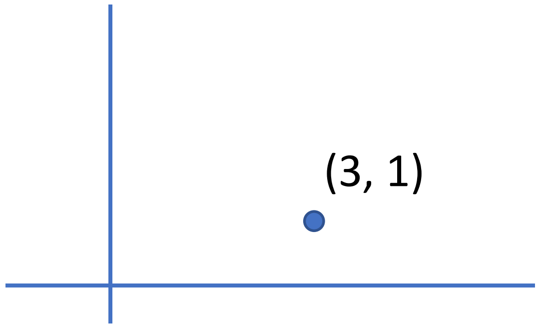
$$\boldsymbol{v} \in \mathbb{R}^n$$

- $\boldsymbol{v}$ a vector with n real entries/elements/components
- n is the size/length of the vector
- $\boldsymbol{v}$ is denoted as follows:

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

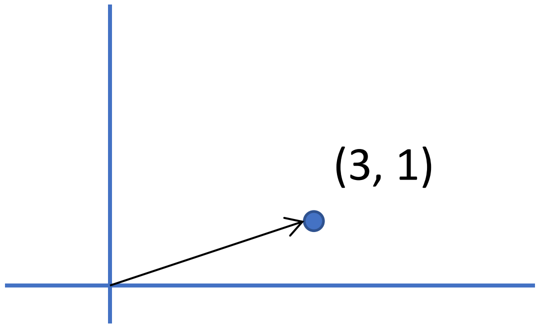
Geometric Interpretation of Vectors

- A vector can be interpreted geometrically as: a point in space or a geometric vector (an object with direction and magnitude).



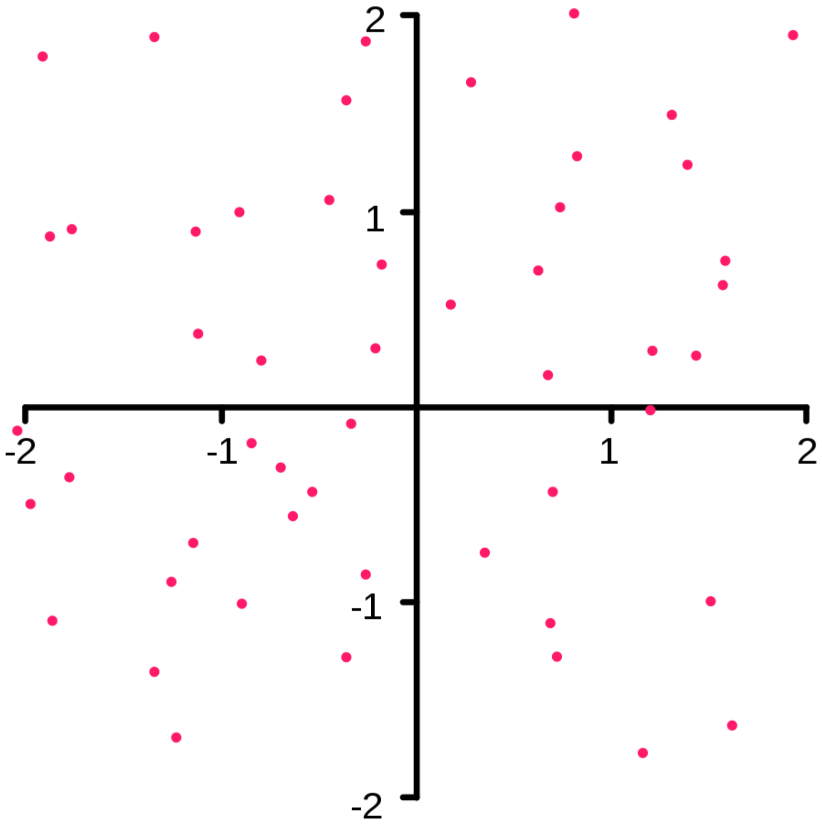
1. Point

OR



2. Geometric vector:
object with direction and magnitude

Geometry of Vectors: Points



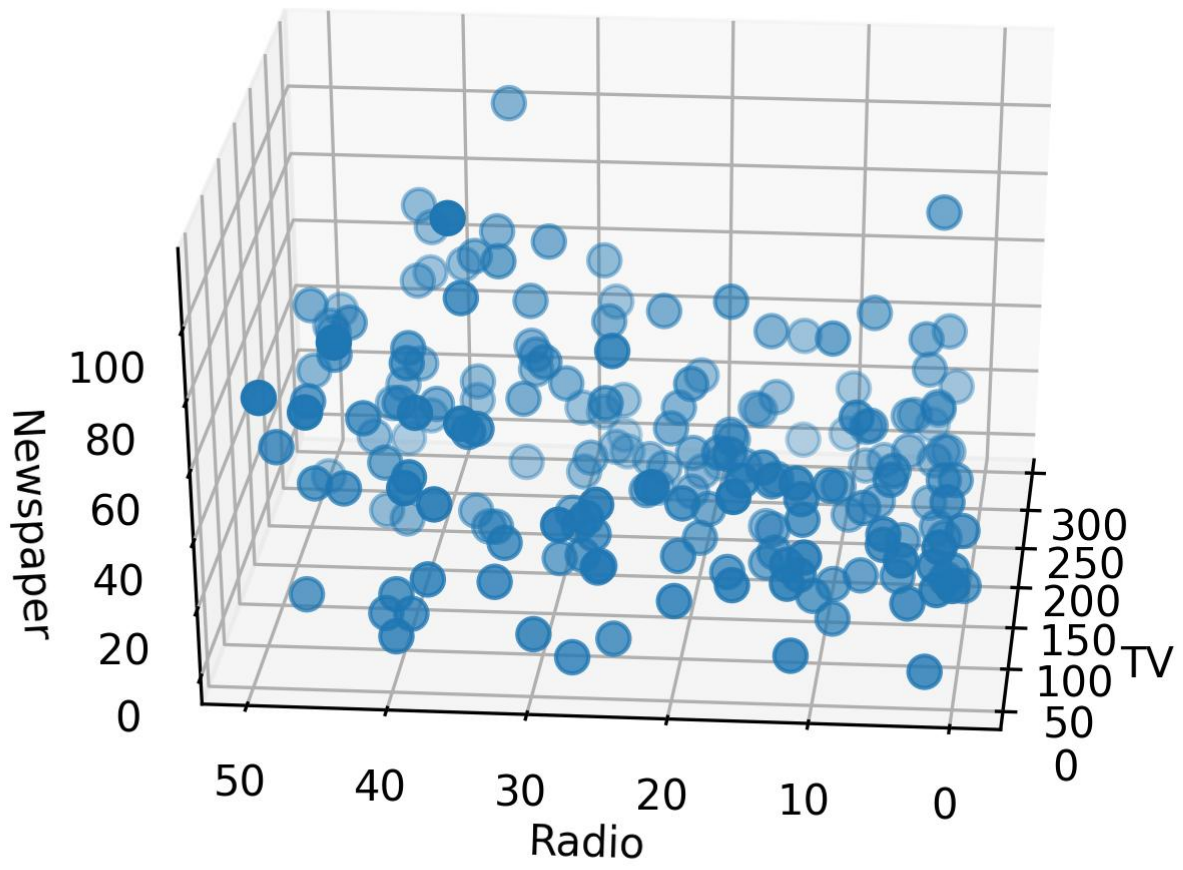
- The components of a vector of size m can be seen as the coordinates of a point in the m-dimensional space.
- The way we can distinguish one point from another is by ***their positions*** in the space.
- Examples of 2D points:
  - (0.75, 0.2)
  - (1.25, 0)
  - (1.25, -1.75)

Examples of 2D points

- Latitude and longitude form a 2D coordinate system or space
- Used to describe locations across the globe
- MSOE: (43.041070, -87.909420)
- Miller Park: (43.011790, -87.967780)
- Margate, FL: (26.242530, -80.204920)

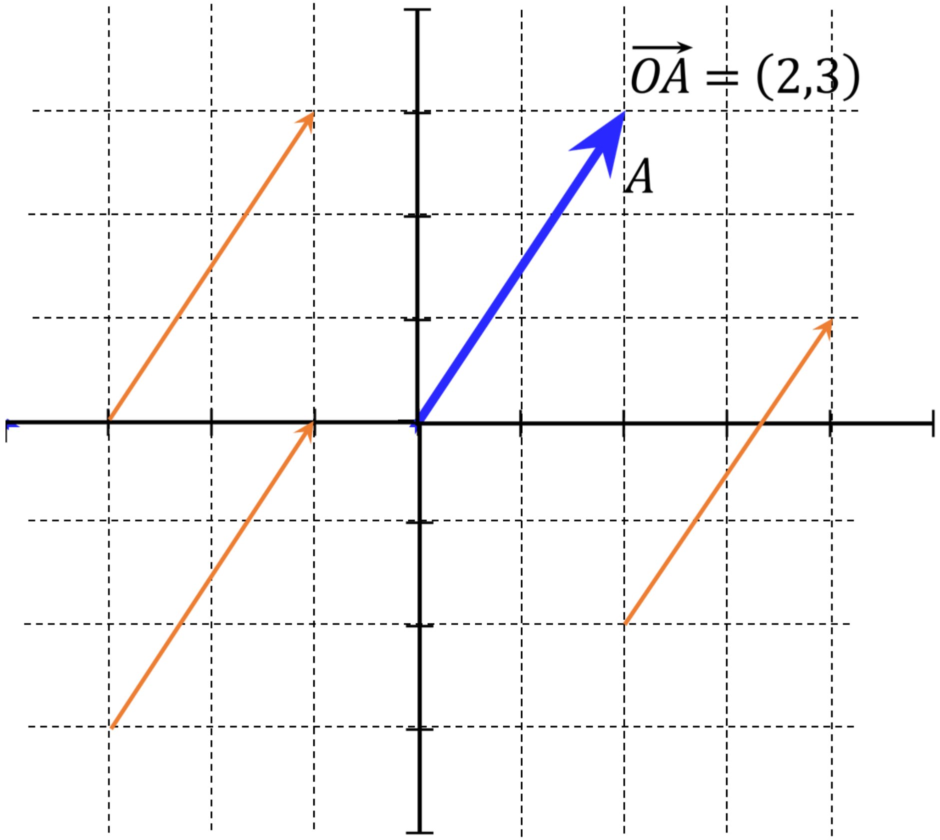
https://www.ergosum.co/cooking-across-latitudes/

Example of 3D Points



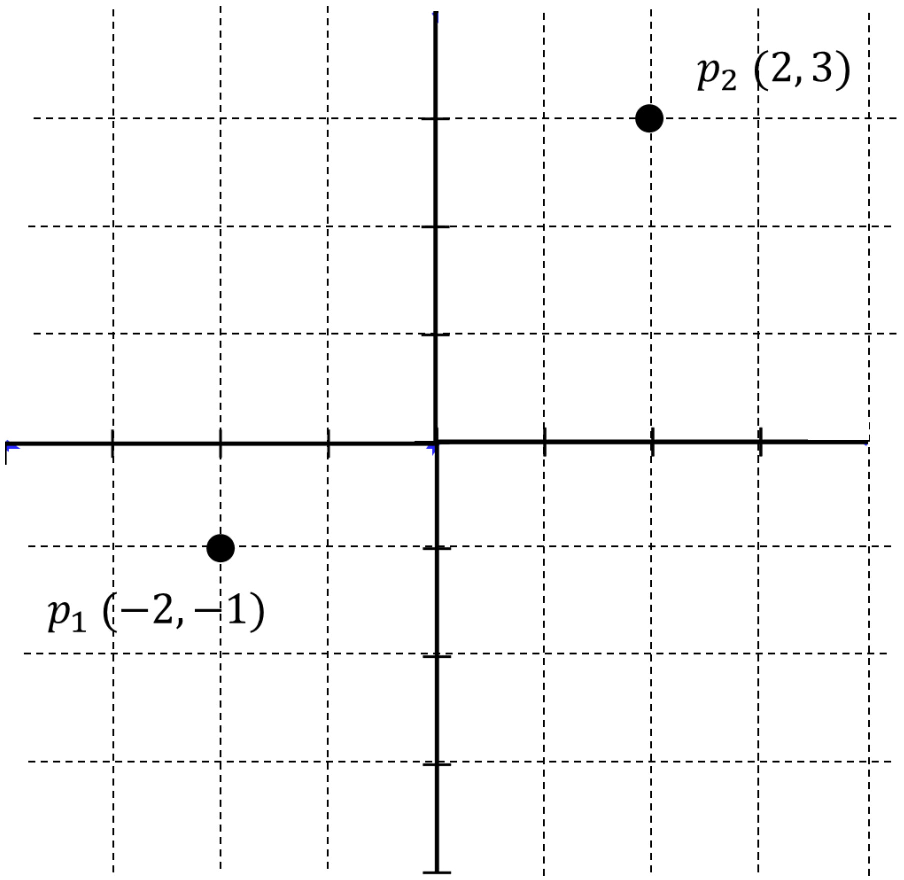
- Each point is the client of an advertising company
- The coordinates refer to the number advertisements they bought
  - TV advertisements
  - radio advertisements
  - newspaper advertisements

Geometry of Vectors: Geometric Vector



- A m-vector can be also interpreted as an object that has ***magnitude and direction***: geometric vector.
- It can be depicted as a directed line segment going from one point to another, where the components of the vector tell us in which direction to go and how far.
- To geometrically represent the vector (2, 3) , we can start from any point and then take 2 steps to the right and 3 steps up. Any of the vector drawn in the left figure can represent the vector (2, 3).
- The way we can distinguish one geometric vector from another is **not by *their positions*** in the space, but by their direction and magnitude.

Creating a Geometric Vector



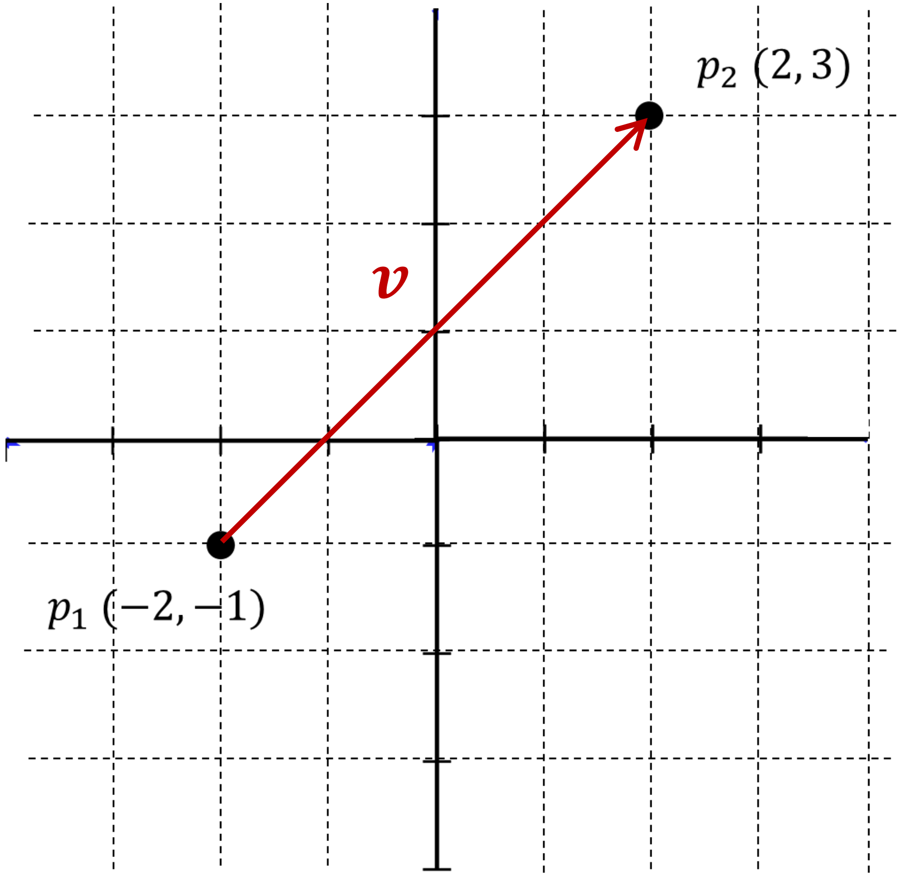
- Given the position of two points, we can compute the components of the geometric vector going from one point to another.

- Example:

$$p_1 = (-2, -1)$$

$$p_2 = (2, 3)$$

Creating a Geometric Vector



p_1 = (-2, -1)

p_2 = (2, 3)

- The components of the vector going from p_1 to p_2 is given by:

$$\boldsymbol{v} = p_2 - p_1 = \begin{bmatrix} 2-(-2) \\ 3-(-1) \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

Vectors

- Note 1:

In machine learning, when we want to interpret a model geometrically, we might treat a given vector as a point or as a geometric vector depending on the application and what the vector represents.

- Note 2:
  - In **linear algebra**, a vector is also given an orientation: it could be a row vector or column vector.
  - The convention followed in most linear algebra references is that the default orientation is column vector.
  - We will try to follow this convention in this class: any vector $\boldsymbol{v}$ is a column vector, unless stated otherwise.

Scalar Multiplication & Vector Addition

- Vectors can be multiplied by a scalar:

If $\boldsymbol{v} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$, then:

$$5\boldsymbol{v} =$$

- Vectors can be added:

If $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$, then:

$$\boldsymbol{u} + \boldsymbol{v} =$$

Dot Product & Norm

- The dot product is a fundamental operation of linear algebra:

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

- If $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$, then:

$$\boldsymbol{u} \cdot \boldsymbol{v} =$$

- The magnitude or norm of a vector can be computed as the square root of the sum of the squares of its elements:

$$\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

- If $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, then $\|\boldsymbol{u}\| =$

- How can we solve for the norm of $\boldsymbol{u}$ using the dot product?

Dot Product & Norm

Note: If $\boldsymbol{u}$ and $\boldsymbol{v}$ are both column vectors, then the dot product $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T \boldsymbol{v}$

- The dot product is a fundamental operation of linear algebra:

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

- If $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$, then:

$$\boldsymbol{u} \cdot \boldsymbol{v} =$$

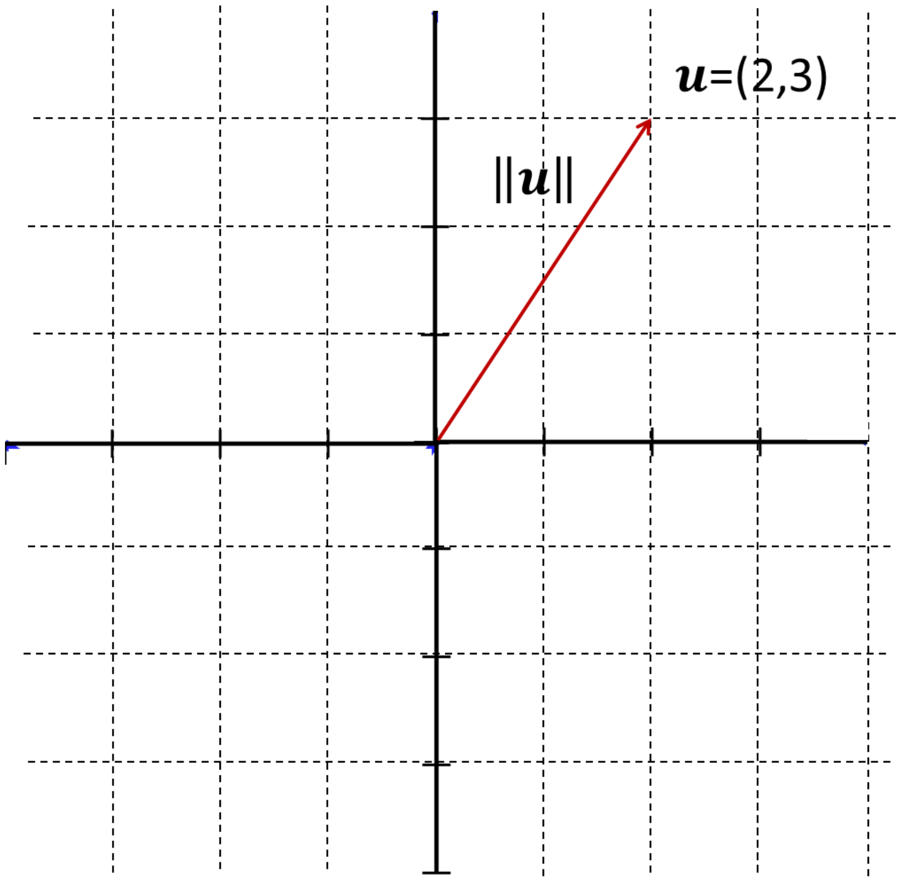
- The magnitude or norm of a vector can be computed as the square root of the sum of the squares of its elements:

$$\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

- If $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, then $\|\boldsymbol{u}\| =$

- How can we solve for the norm of $\boldsymbol{u}$ using the dot product?

Vector Norm



- Geometrically, the norm of a vector is the length of the line segment that represents the vector.
- A unit vector is vector with magnitude of 1:

  If $\boldsymbol{u}$ is a vector, then its unit vector version is:

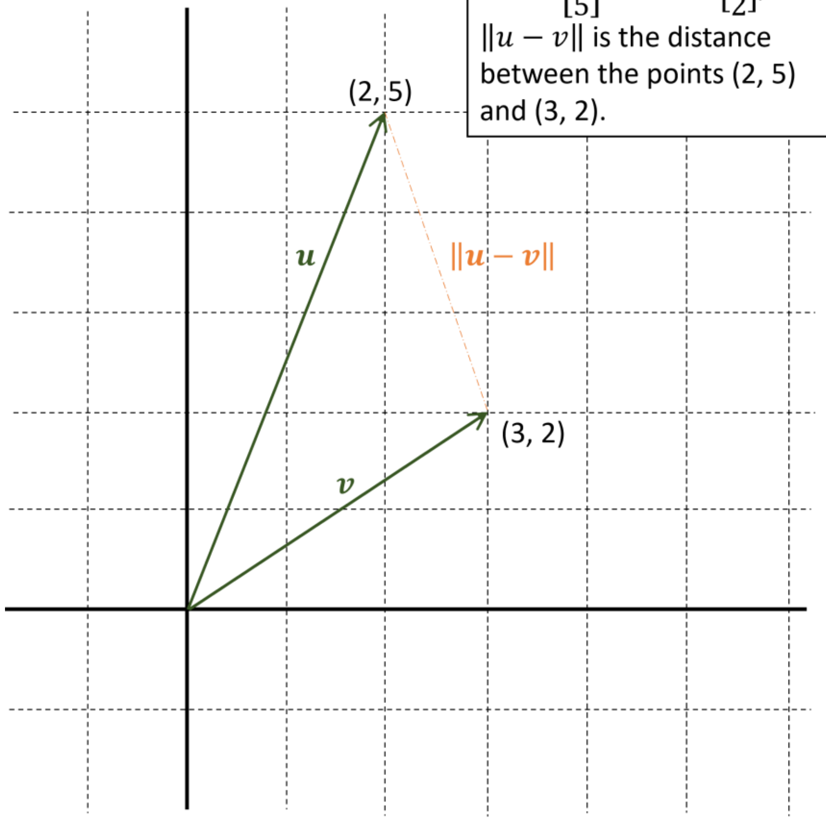
$$\boldsymbol{u}_{unit} =$$

Euclidean Distance

- The Euclidean distance between two vectors u and v is the norm of their difference:

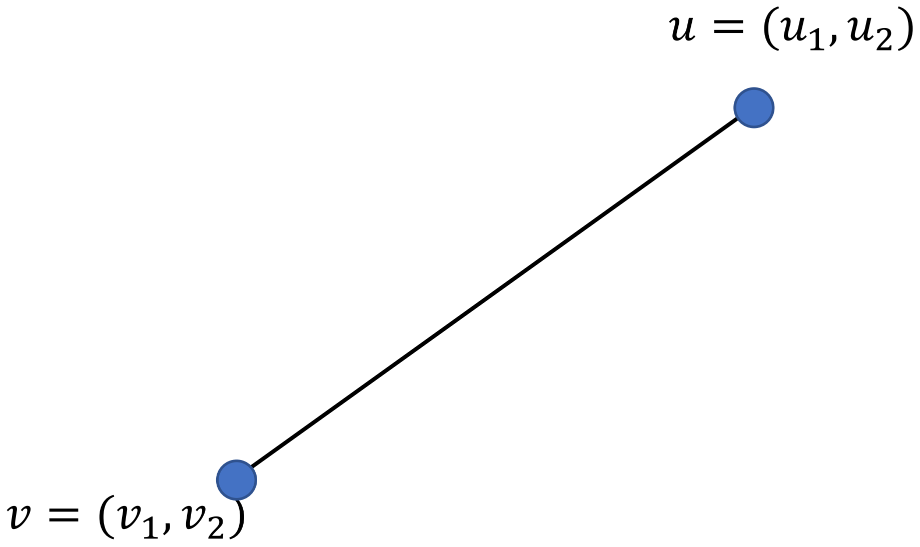
$$\mathrm{dist}(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\| = \sqrt{(u-v)\cdot(u-v)}$$

$$= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$$



If u $= \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and v $= \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, then $\|u - v\|$ is the distance between the points (2, 5) and (3, 2).

Euclidean Distance

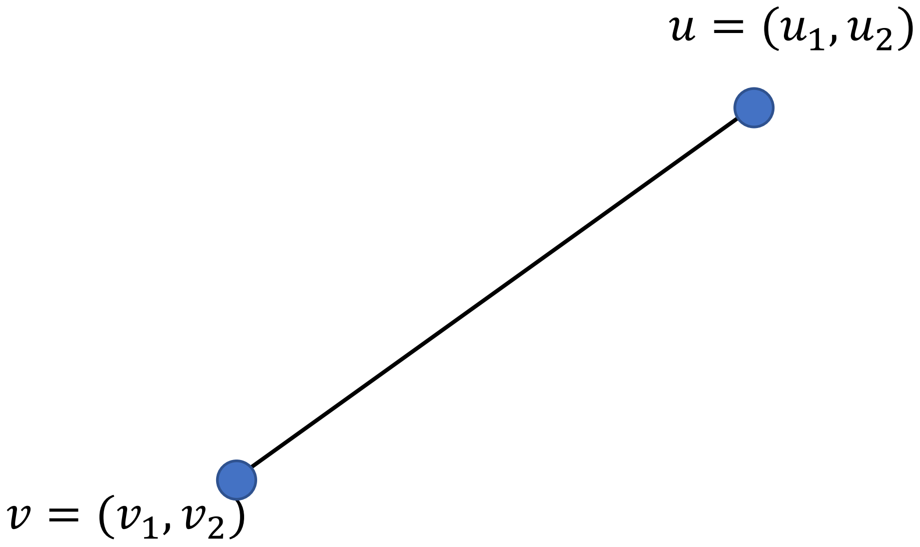


You are probably familiar with Euclidean distance:

- Distance between 2D points (u_1, u_2) and (v_1, v_2):
$$d = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2}$$
- Distance between 3D points (u_1, u_2, u_3) and (v_1, v_2, v_3)
$$d = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}$$

Euclidean distance can be extended to any number of dimensions

Euclidean Distance in terms of norm

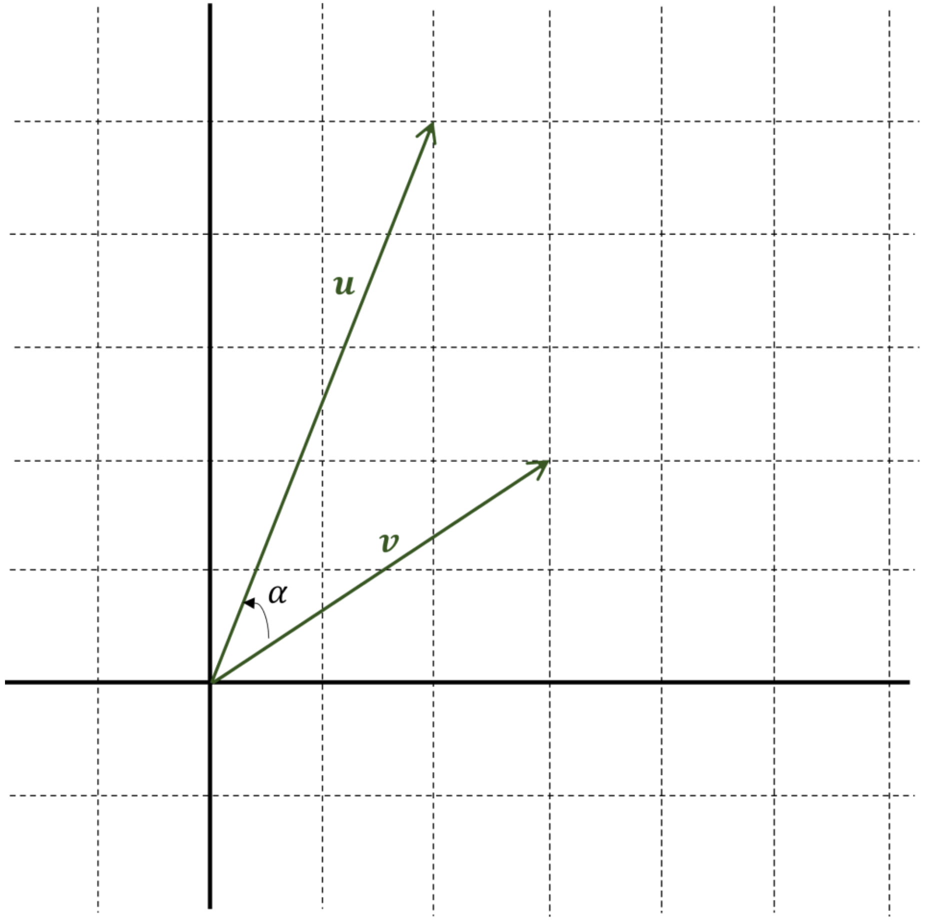


$$d = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2} = \|\boldsymbol{u} - \boldsymbol{v}\|$$

Dot Products & Vector Norms

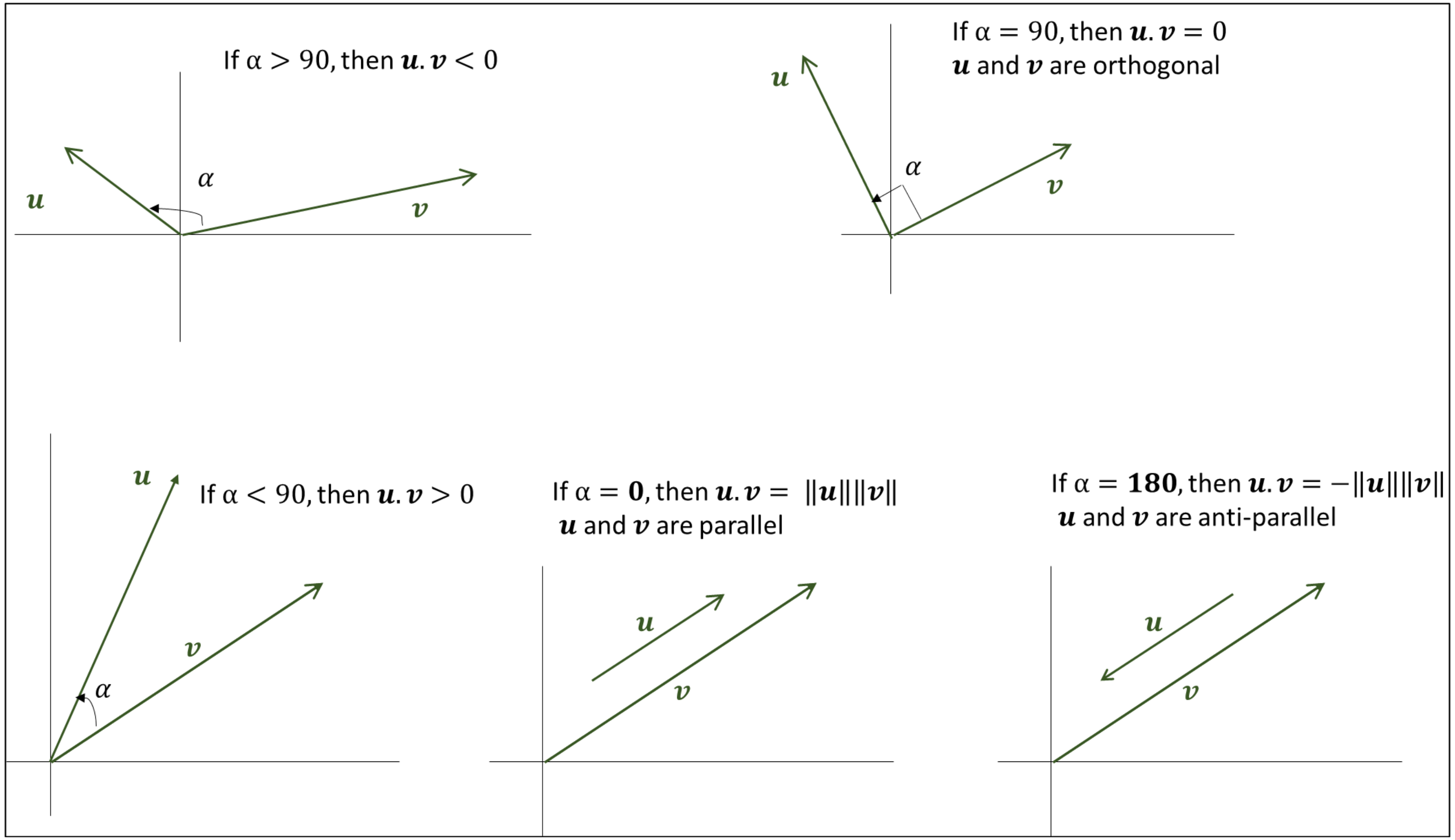
- The dot product of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ can be also given by:

$$\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos \alpha$$

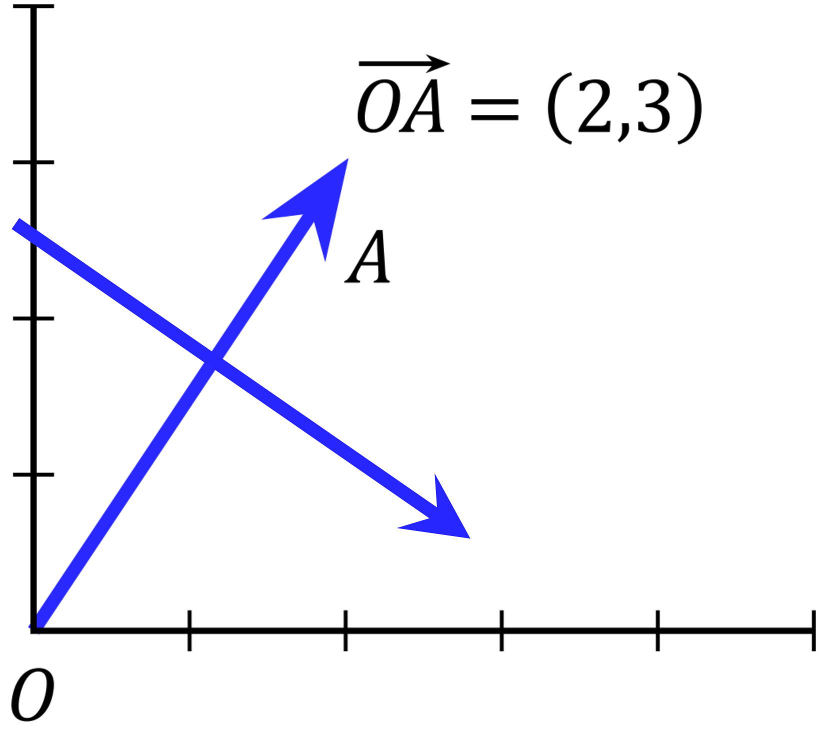


Dot Products & Vector Norms

$\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos \alpha$



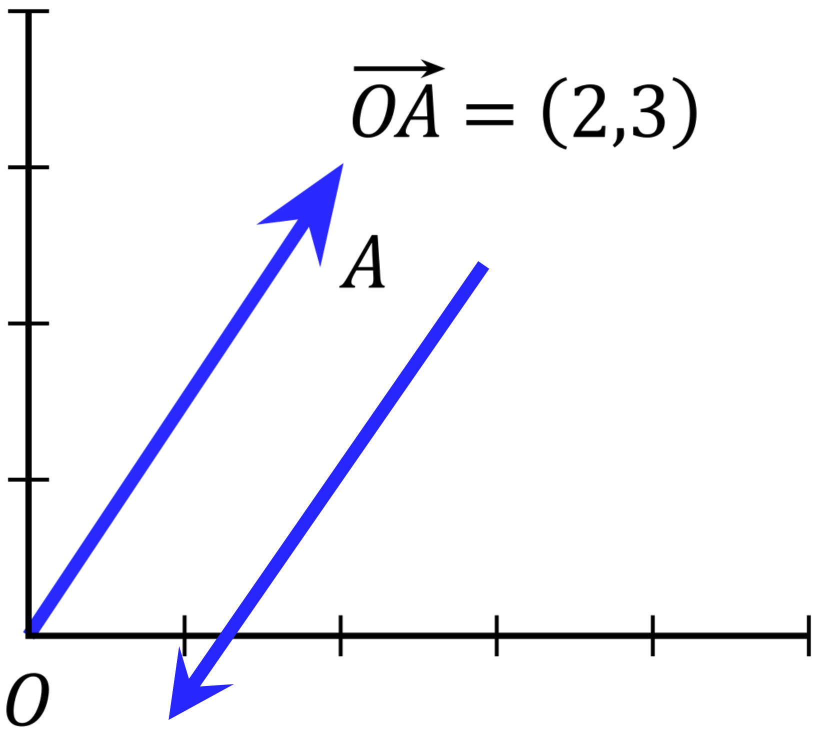
Vectors: Dot Product



If two vectors are perpendicular (orthogonal), then their dot product is 0

$$\vec{u} \cdot \vec{v} = 0$$

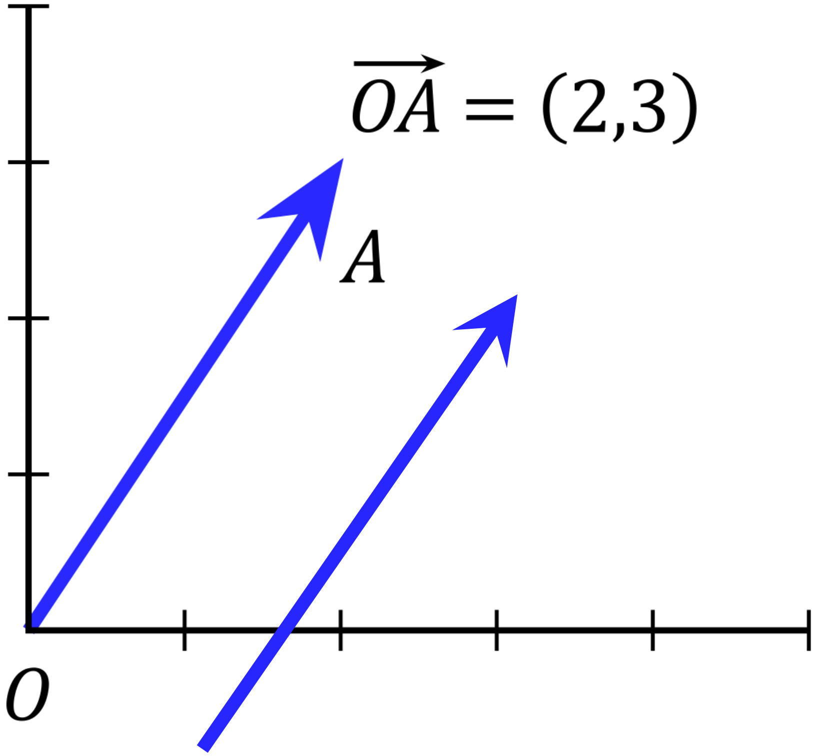
Vectors: Dot Product



If two vectors are anti-parallel, then the dot product of their unit vectors is -1

$$\frac{\vec{u}}{||\vec{u}||} \cdot \frac{\vec{v}}{||\vec{v}||} = -1$$

Vectors: Dot Product



If two vectors are parallel, then the dot product of their unit vectors is 1

$$\frac{\vec{u}}{||\vec{u}||} \cdot \frac{\vec{v}}{||\vec{v}||} = 1$$

Matrices

Matrices

- Matrices are an ordered collection of vectors (either column vectors or row vectors).
- Matrices are used to store and perform operations on groups of vectors in one operation.
- A matrix is a rectangular array of numbers written between squared brackets:

$$\begin{bmatrix} 1 & 3 \\ 5 & 7 \\ 9 & 11 \end{bmatrix}$$

Vectors

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

- The matrix contain real entries organized into:
  - n rows
  - m columns

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix}$$

Vectors

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

- The matrix contain real entries organized into:
  - n rows
  - m columns

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix}$$

Matrix elements can be described using indexing like so: $a_{i,j}$. The index i refers to the row, while the index j refers to the column.

Matrices

- **Square matrices**: A square matrix is a matrix that has an equal number of rows and columns.
- **Matrix transpose**: If A is an $n \times m$ matrix, its transpose, denoted A^T , is the $m \times n$ matrix, where each row in A becomes a column in A^T. For example,

$$\begin{bmatrix} 1 & 3 \\ 5 & 7 \\ 9 & 11 \end{bmatrix}^T =$$

Matrix-Vector Multiplication

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

$$\begin{bmatrix} 1 & 3 \\ 5 & 7 \\ 9 & 11 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

- Multiplying a matrix and a vector produces a new vector.
- To get each element of the new vector:
  - each row of the matrix is dotted with vector
- The number of elements of **b** needs to be equal to the number of columns of A.

Matrix-Matrix Multiplication

$$AB = C$$

$$\begin{bmatrix} 1 & 3 \\ 5 & 7 \\ 9 & 11 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ & 2 \\ -2 & \end{bmatrix}$$

- We can also multiple two matrices.
- To get every element in the output matrix:
  - we dot each row of the first matrix with every column of the second matrix.

The number of columns in the first matrix must match the number of rows in the second matrix (The dimensions of the two matrices must be compatible).

Matrix-Matrix Multiplication

$$AB = C$$

$$\begin{bmatrix} 1 & 3 \\ 5 & 7 \\ 9 & 11 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ & 2 \\ -2 & \end{bmatrix}$$

- We can also multiple two matrices.
- To get every element in the output matrix:
    - we dot each row of the first matrix with every column of the second matrix.

Geometric Example of Matrix-Vector Multiplication

- Matrix-vector multiplication can be seen as vector transformation known as linear mapping.

$$\boldsymbol{v} = A\boldsymbol{u}: \quad \boldsymbol{u} \xrightarrow{A} \boldsymbol{v}$$

- We say A transforms the vector $\boldsymbol{u}$ into a new vector $\boldsymbol{v}$.
- The type/meaning of transformation depends on how we choose A.

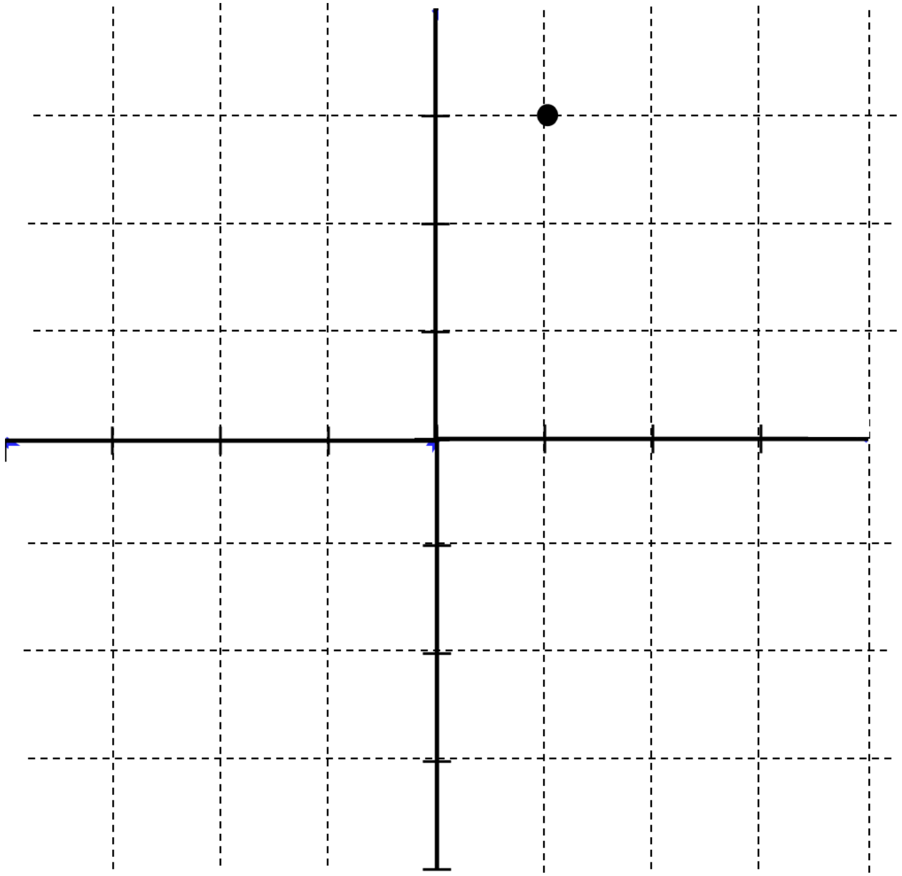
Example 1 - Negation

- Suppose,

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } \boldsymbol{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

- Assume **u** is a point in the 2-dimensional space. How would **u** change after multiplying it by A?

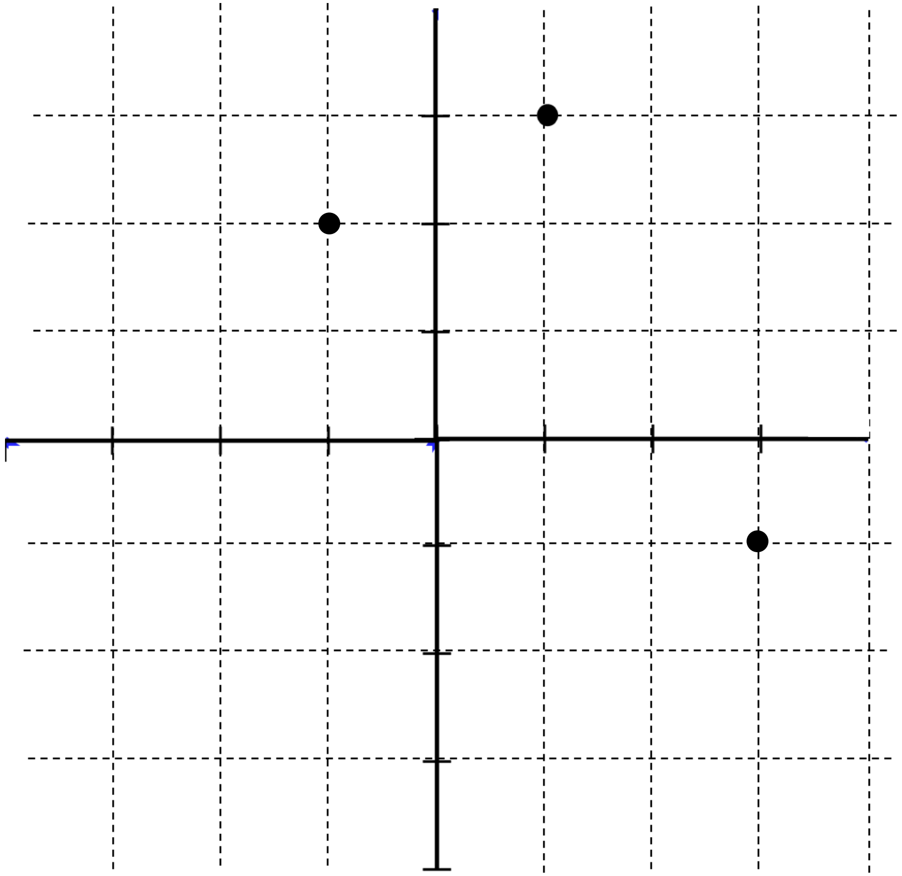
$$A\boldsymbol{u} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} \quad \end{bmatrix}$$



Example 1 - Negation

- Now we want to apply the same transformation to two other vectors:

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$



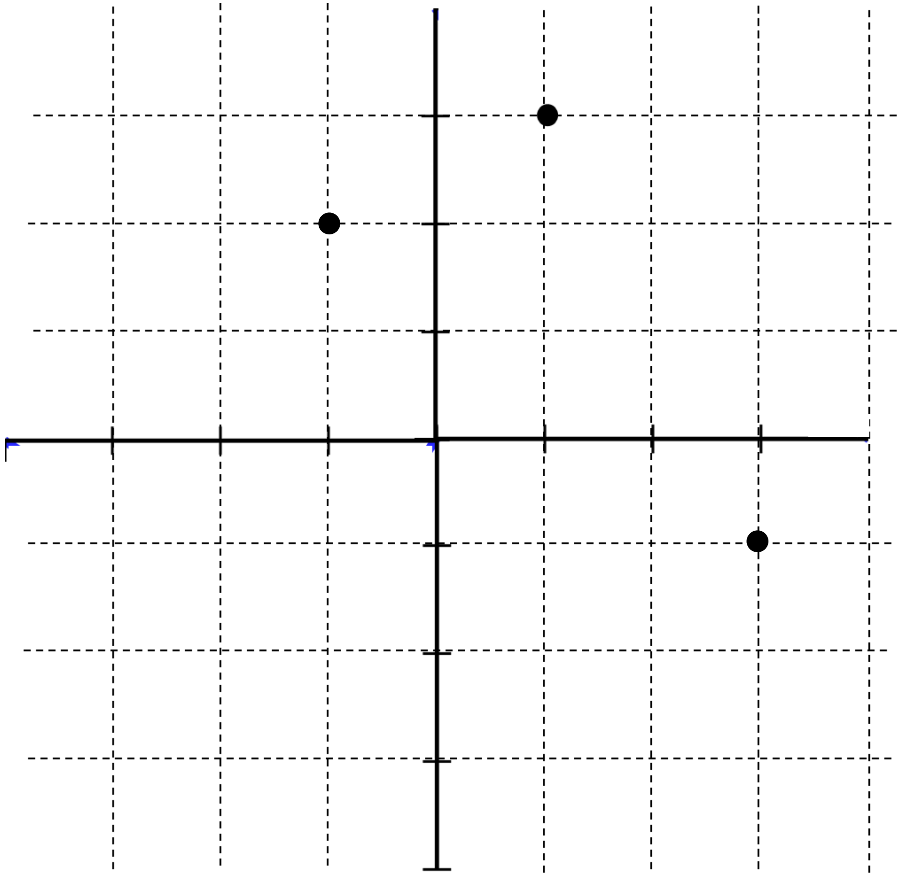
Example 1 - Negation

- Now we want to apply the same transformation to two other vectors:

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

- We can apply the transformation to each vector individually or we can collect the vectors that we want to transform into one matrix to represent a collective transformation:

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



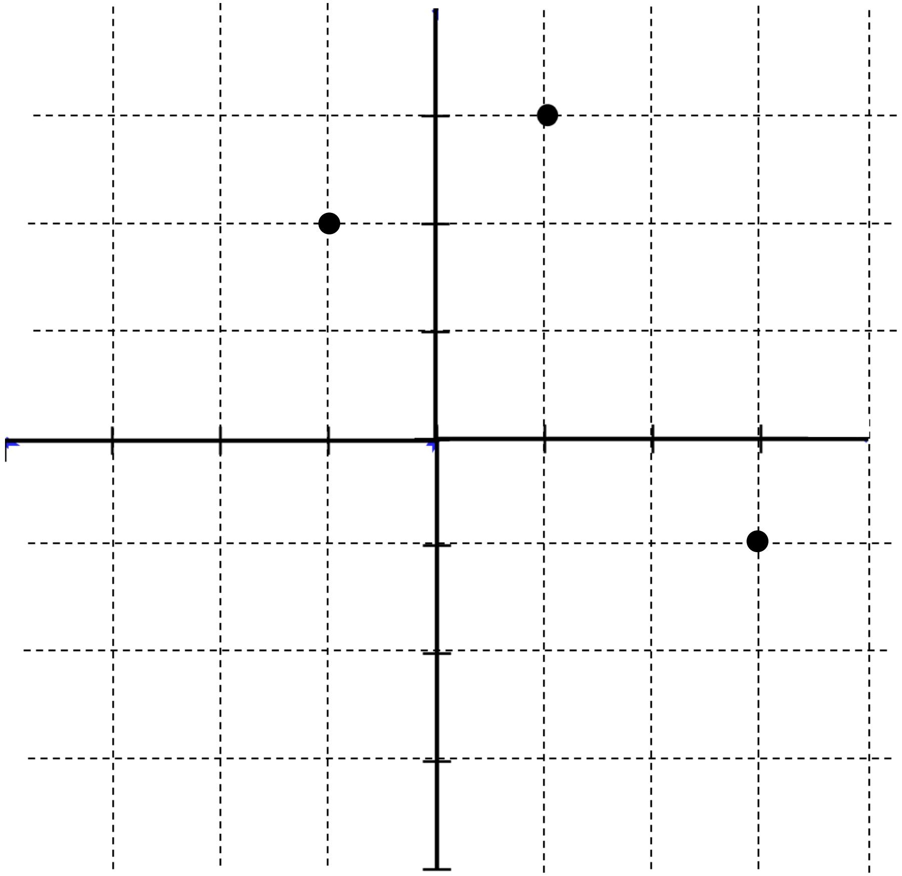
Example 2 - Permutation

- Now we want to apply another transformation to two other vectors:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

- How would the points change in this case?

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

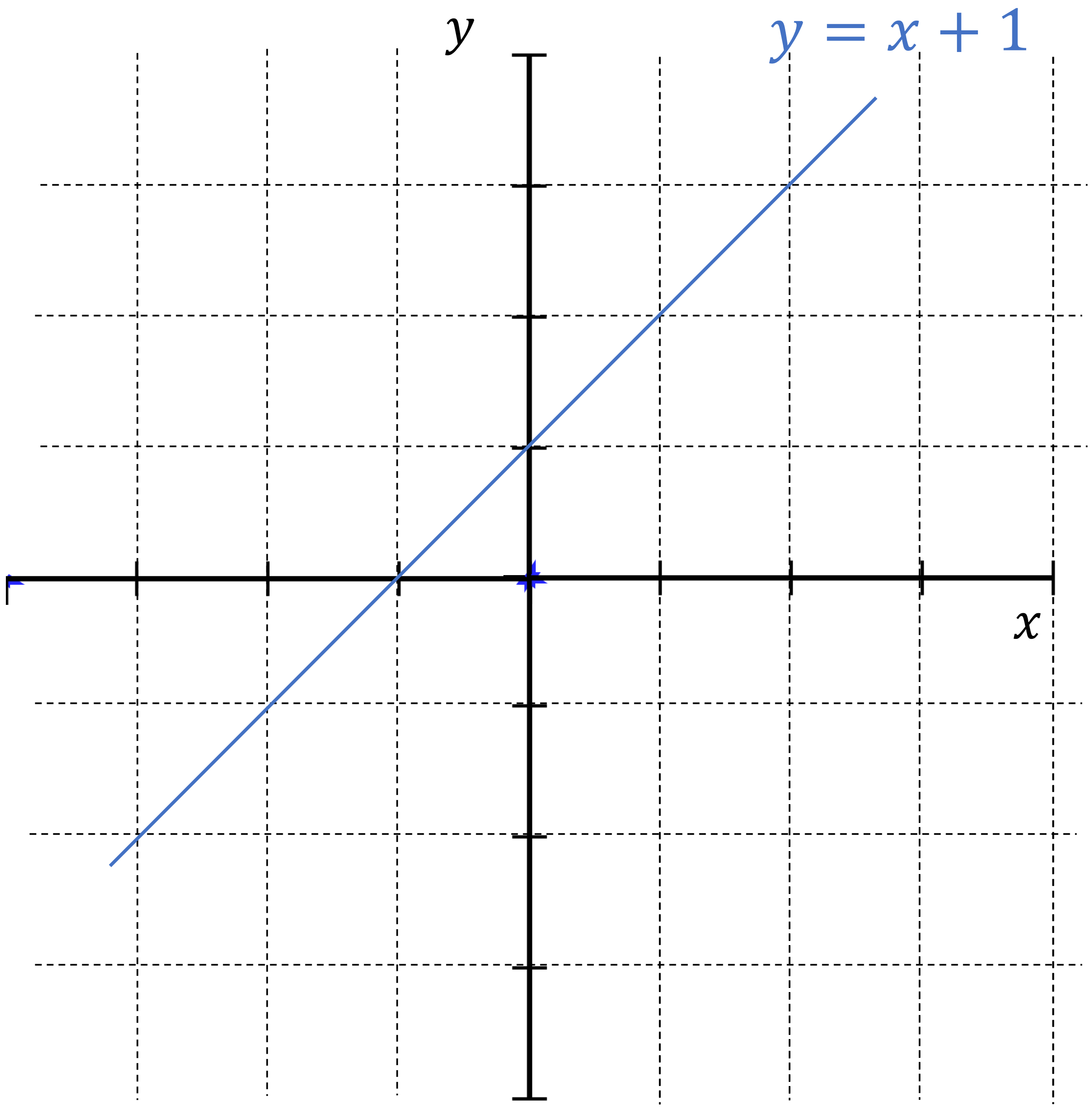


Lines, Planes, Hyperplanes

Line’s Equation

You’re familiar with the following form of a line’s equation:

$$y = x + 1$$

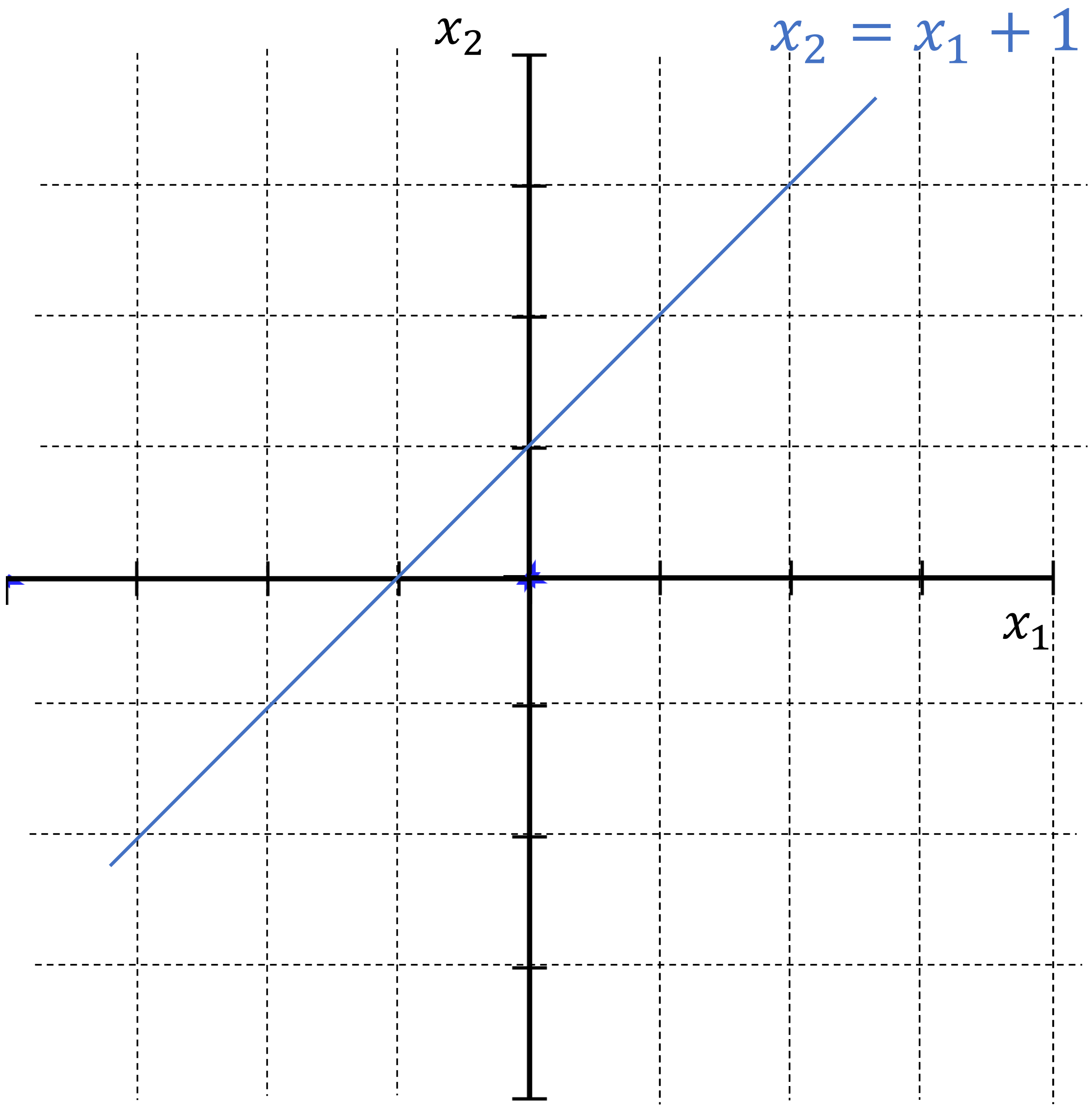


Line's Equation

- In this class (when we're going to talk about linear classification), we might need to label the horizontal axes differently:

$$x_2 = x_1 + 1$$

- We want to you to get used to using variables other than y and x when defining a line's equation.



Line's Equation

- If we re-arrange the terms of the following line's equation:

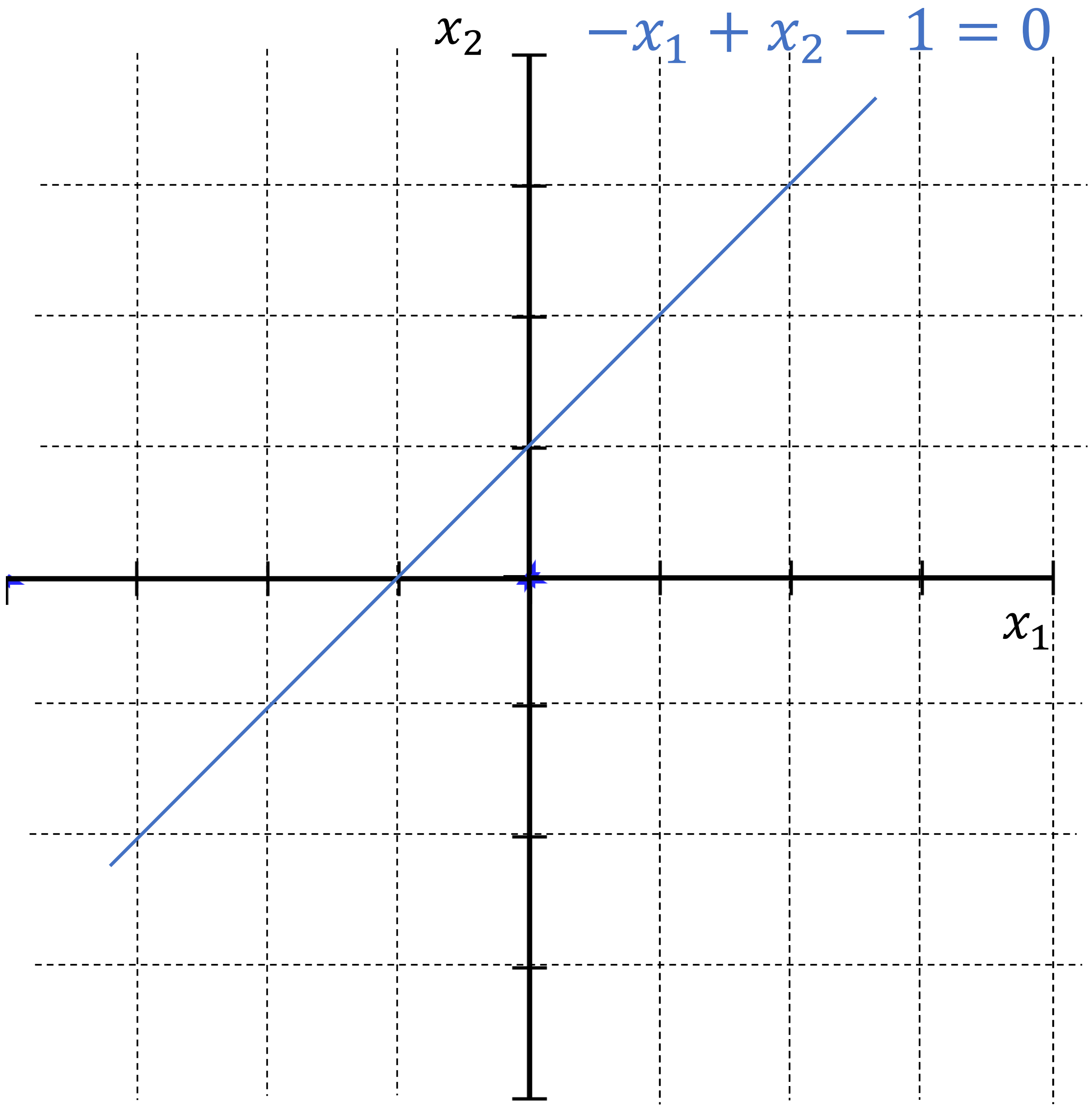
$$x_2 = x_1 + 1$$

we get the standard form of the line's equation:

$$-x_1 + x_2 - 1 = 0$$

- More generally speaking, a line's equation in a 2-dimensional space is given by:

$$w_1 x_1 + w_2 x_2 + w_0 = 0$$



Line's Equation

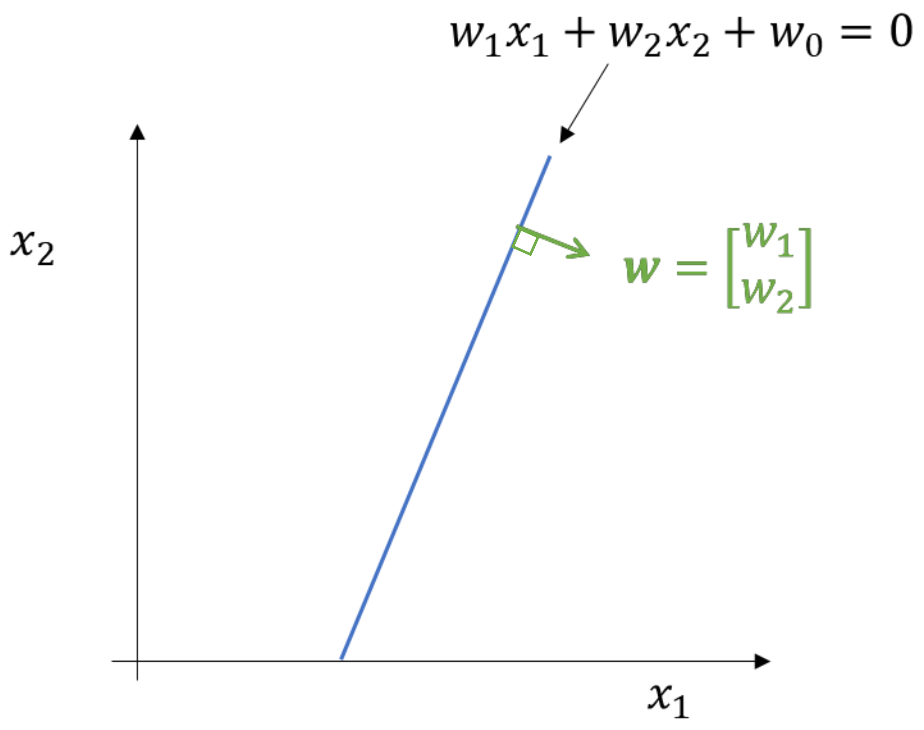
- The standard form of a line's equation is given by:

$$w_1x_1 + w_2x_2 + w_0 = 0$$

- The coefficients w_1 and w_2 represent the components of the line's normal vector:

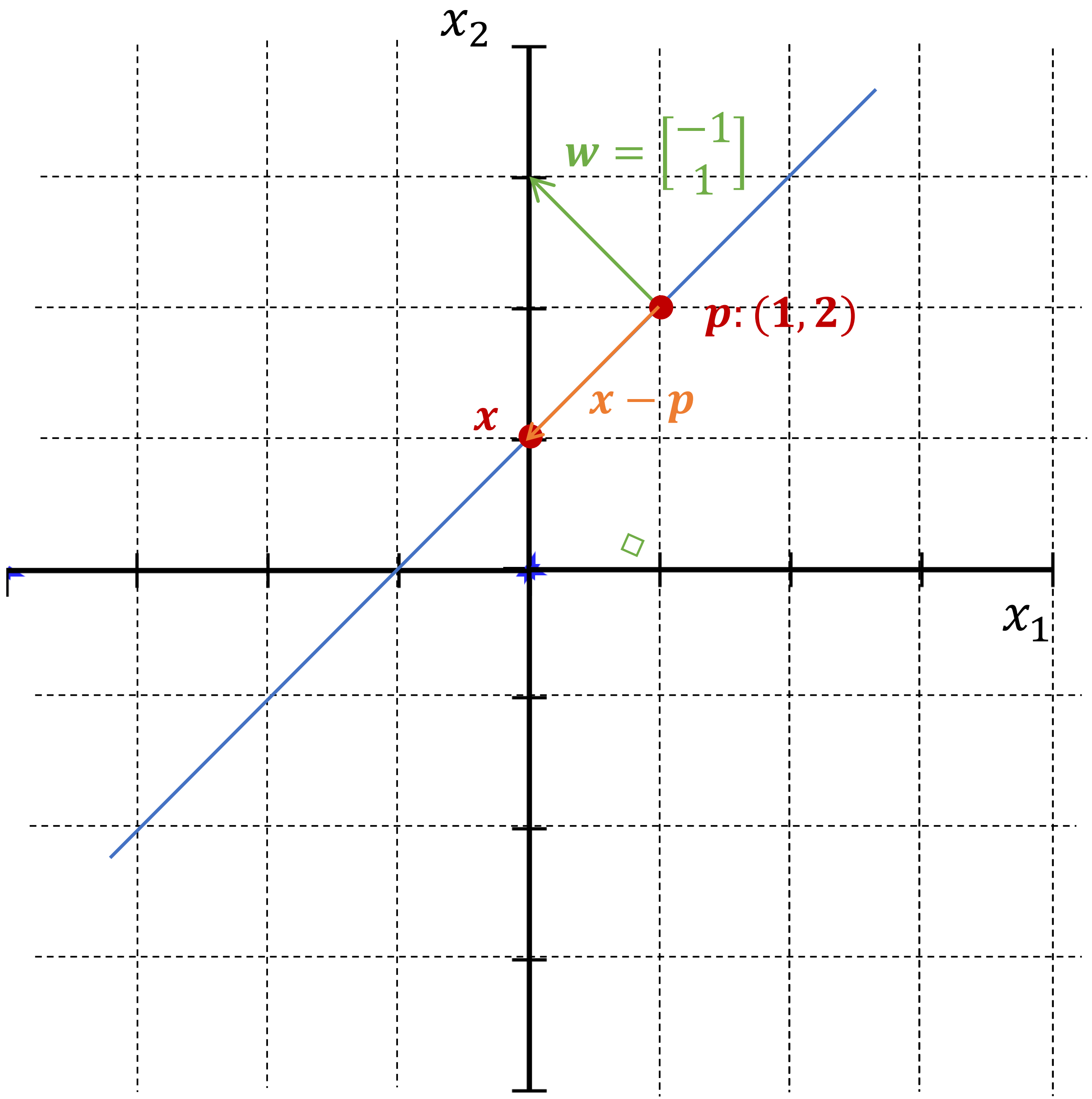
$$\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

- The normal vector is the vector that is orthogonal to the line, and it is what determines the orientation of the line.



Line's Equation

- A line can be determined by its standard equation.
- Or it can be determined using a line's normal vector $\boldsymbol{w}$ and a particular point $\boldsymbol{p}$ that the line goes through.



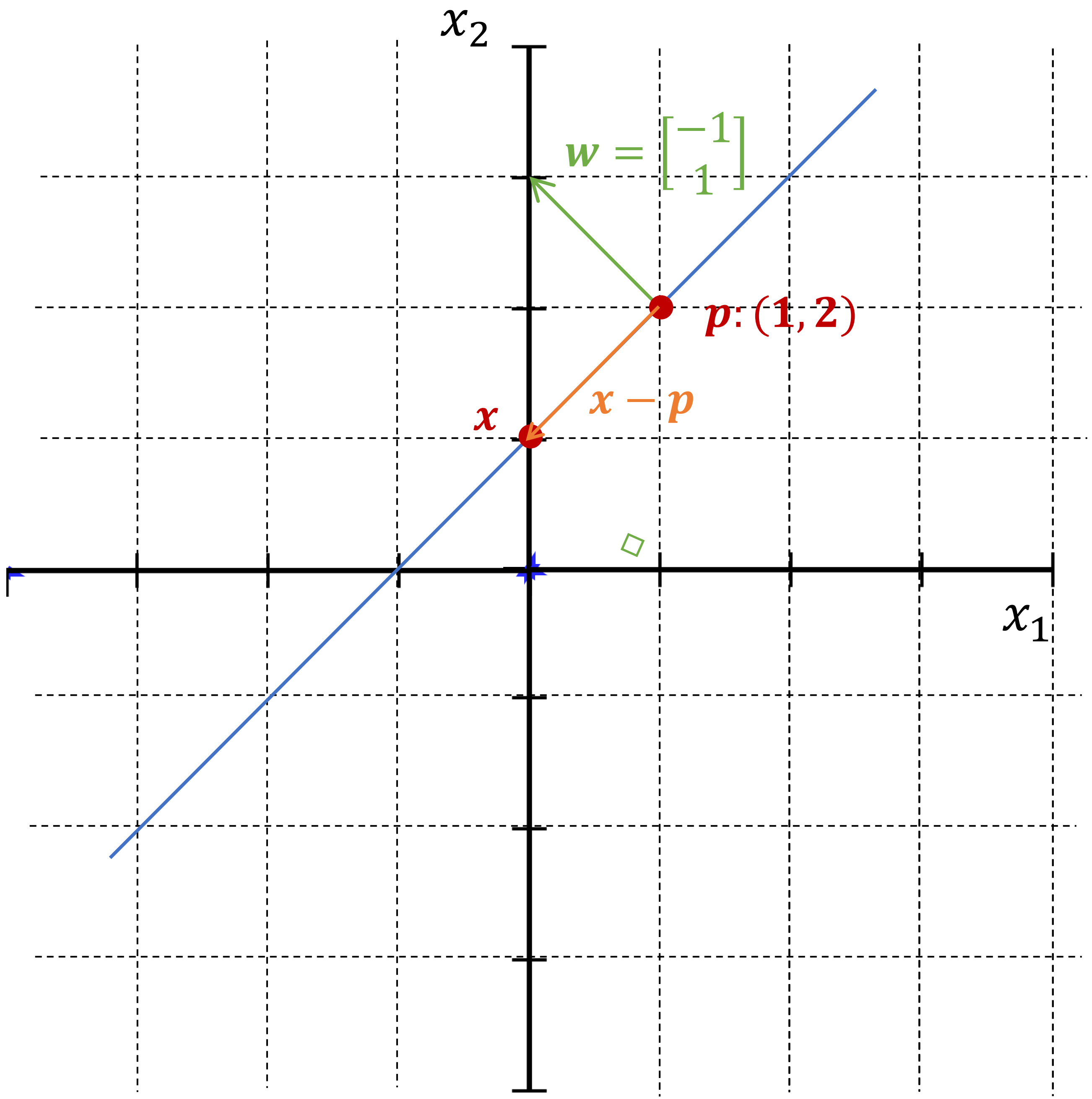
Line's Equation

- Example: determine the equation of the line that goes through $\boldsymbol{p}$=(1,2) and has $\boldsymbol{w} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ a normal vector.

- Let $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ represent a point on the line.

  Then,

$$\boldsymbol{w} \cdot (\boldsymbol{x} - \boldsymbol{p}) = 0$$



Line's Equation

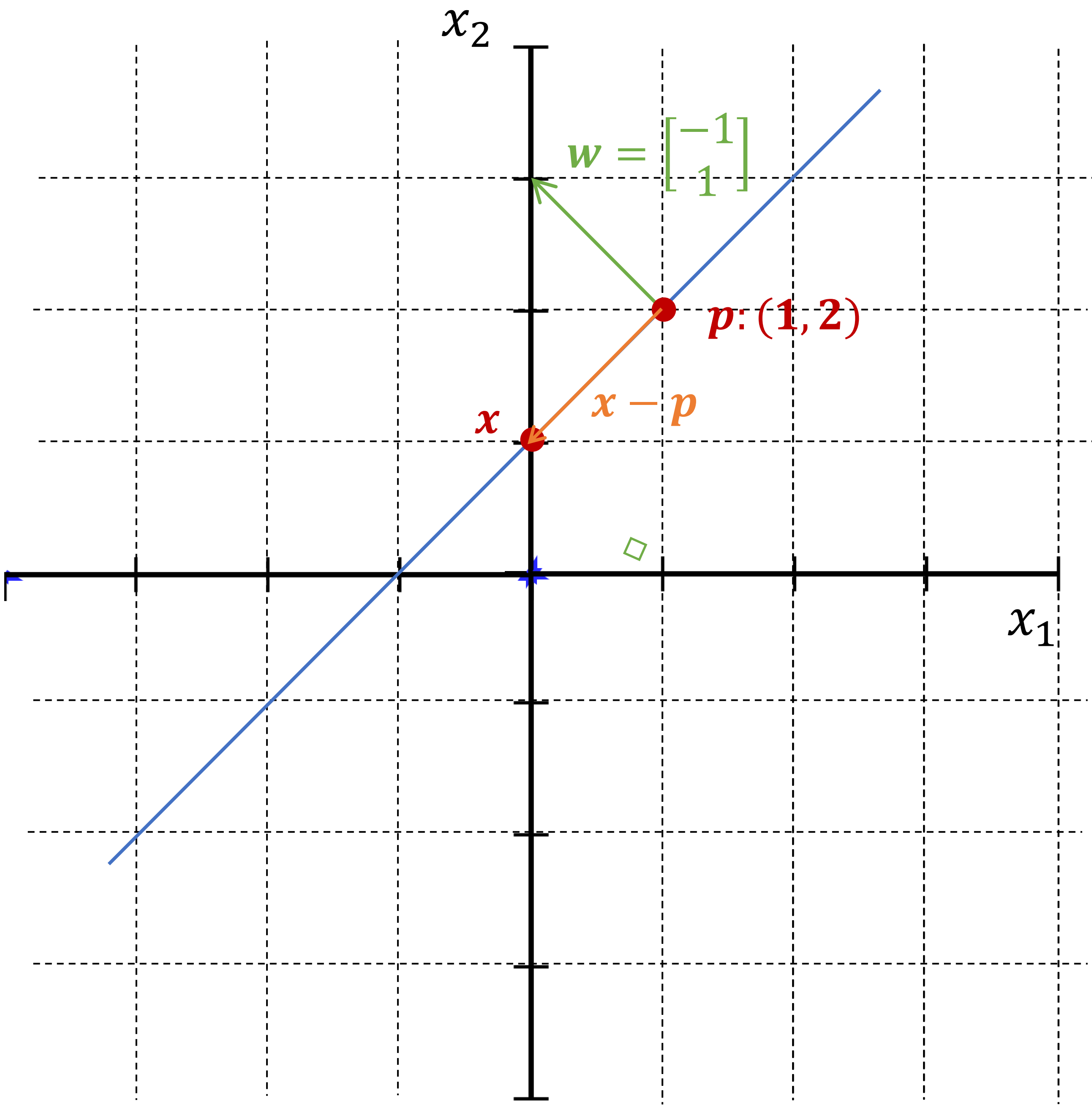
- Example: determine the equation of the line that goes through $\boldsymbol{p}$=(1,2) and has $\boldsymbol{w} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ a normal vector.
- Let $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ represent a point on the line.

Then,

$$\boldsymbol{w} \cdot (\boldsymbol{x} - \boldsymbol{p}) = 0$$

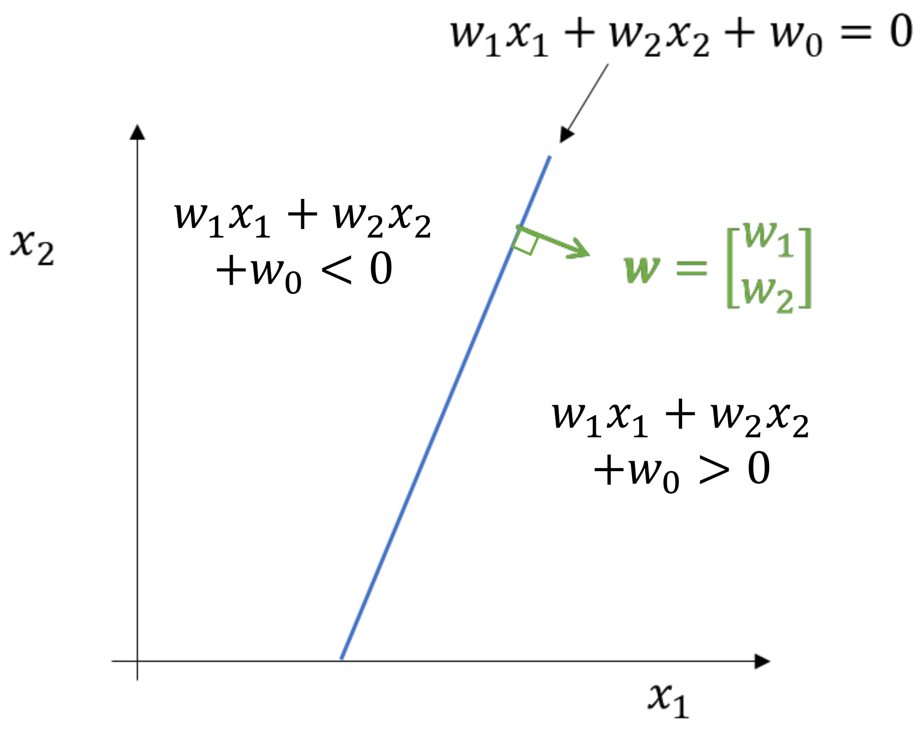
$$w_1(x_1 - 1) + w_2(x_2 - 2) = 0$$

$$-(x_1 - 1) + (x_2 - 2) = 0$$
$$\Rightarrow -x_1 + x_2 - 1 = 0$$



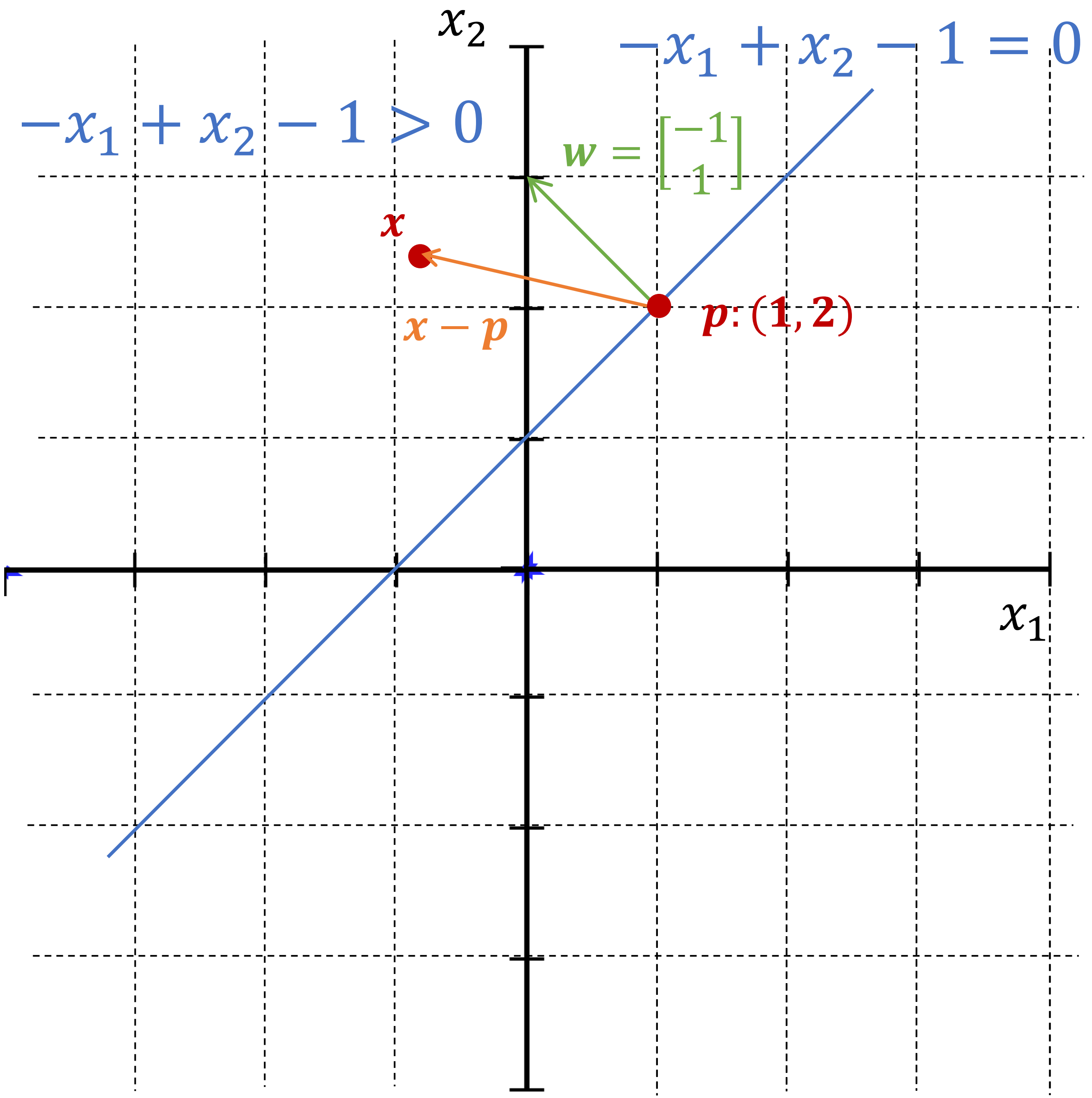
Line's Equation

- Any point $\boldsymbol{x}$ that lies on the line satisfies:
$$w_1x_1 + w_2x_2 + w_0 = 0$$
- Any point $\boldsymbol{x}$ that lies on the same side of the normal vector satisfies:
$$w_1x_1 + w_2x_2 + w_0 > 0$$
- Any point $\boldsymbol{x}$ that lies on the different side of the normal vector satisfies:
$$w_1x_1 + w_2x_2 + w_0 < 0$$
- We say that the line divides the space into half-planes (half-spaces).



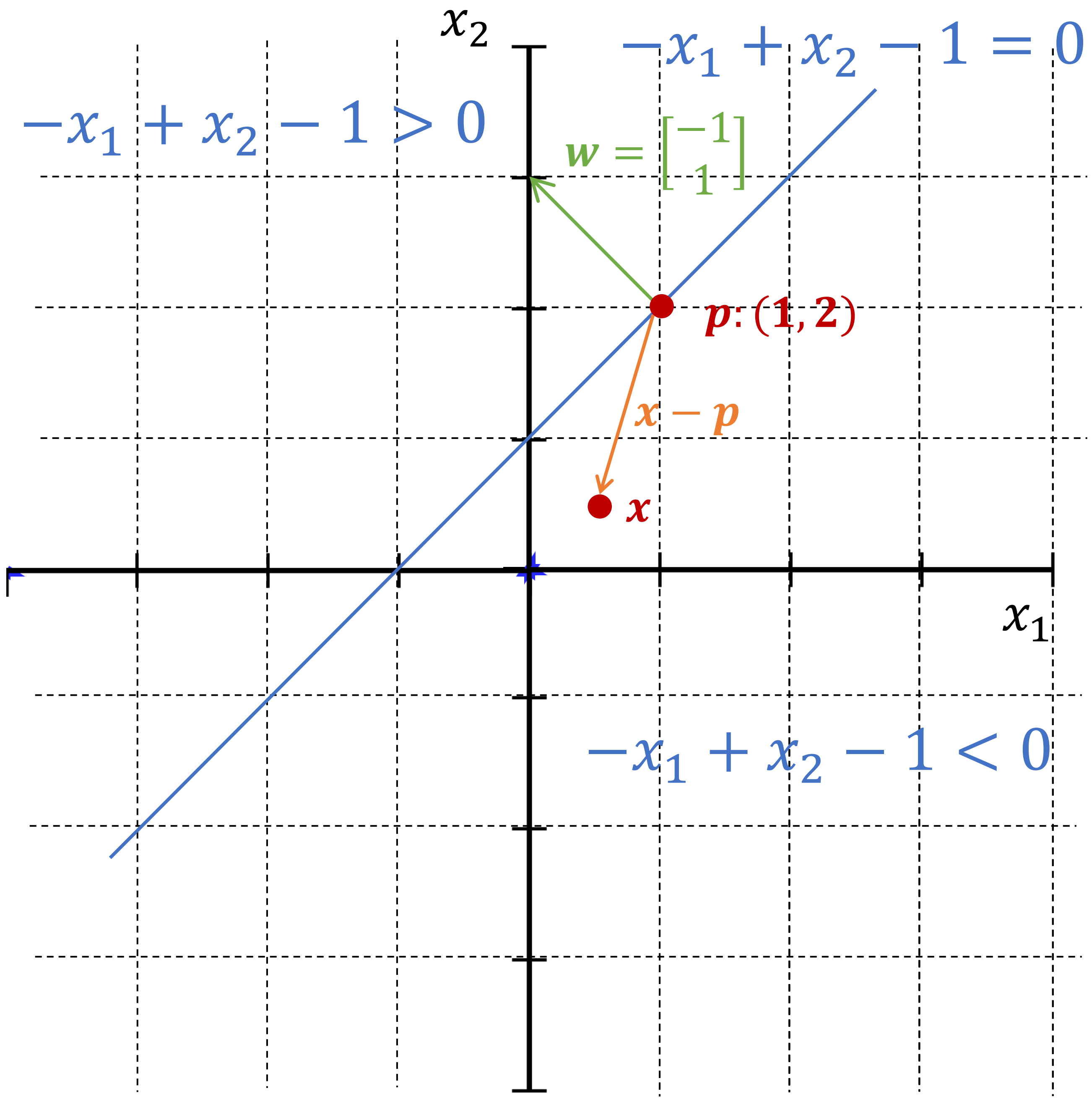
Line's Equation

- Given the coordinates of a point, we can plug in the values using the line's equation to see if the value is positive or negative.
- The faster way is to check whether the point lies on the same/different side of the normal vector.

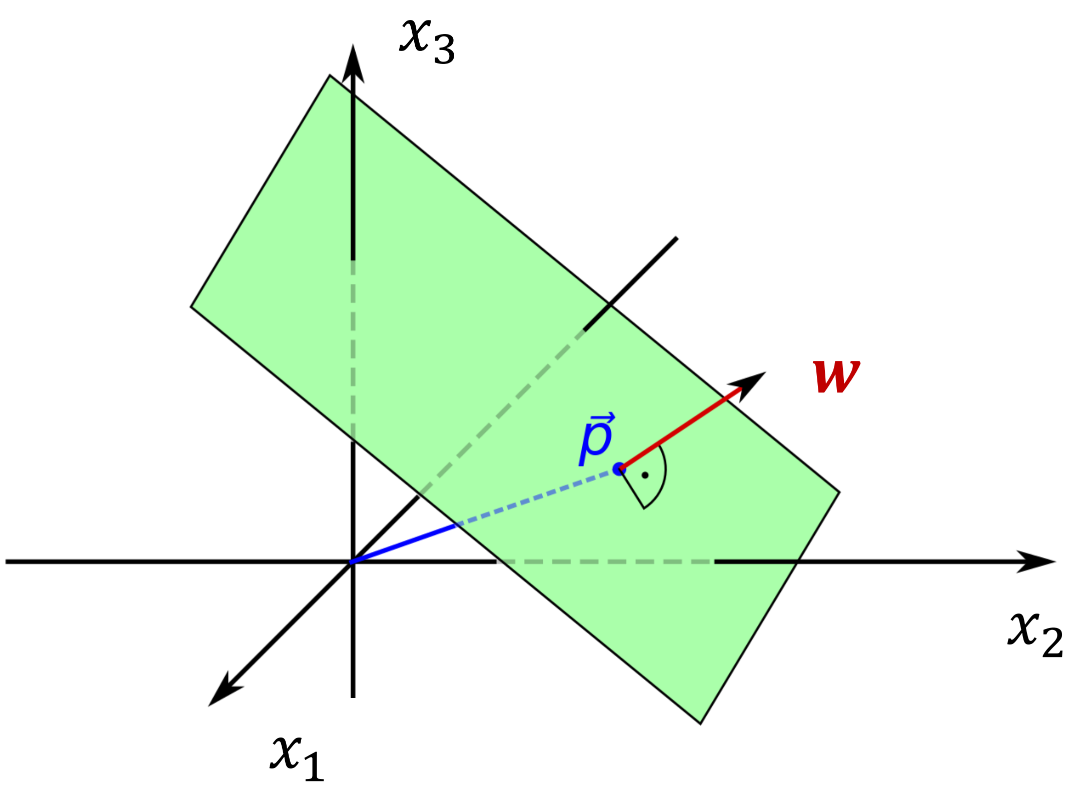


Line's Equation

- The positive and negative values can be seen due to whether the angle between $\boldsymbol{w}$ and $(\boldsymbol{x} - \boldsymbol{p})$ is less then 90 or greater than 90.



Planes

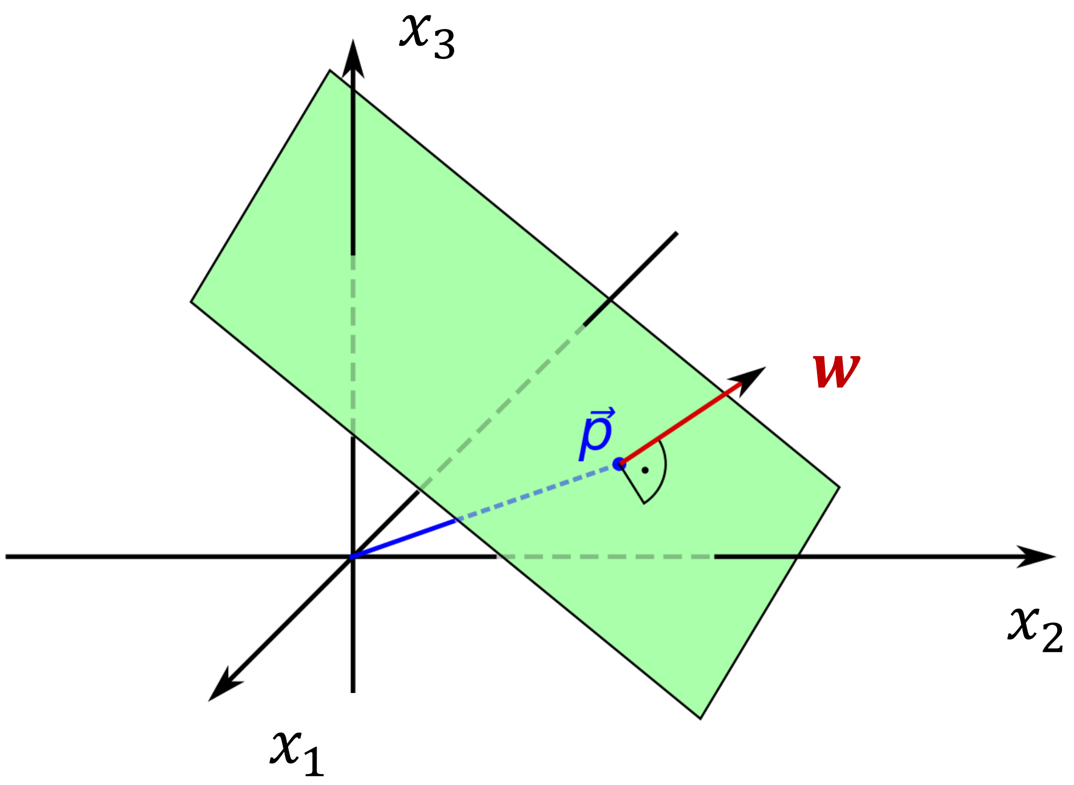


- The generalization of the line to the 3-dimensional space is called the plane and is given by:

$$w_1x_1 + w_2x_2 + w_3x_3 + w_0 = 0$$

- w is the plane's normal vector: it is perpendicular to the surface, and it defines the orientation of the plane.
- A plane is a flat geometric object that divides a space in two half-spaces.

Planes

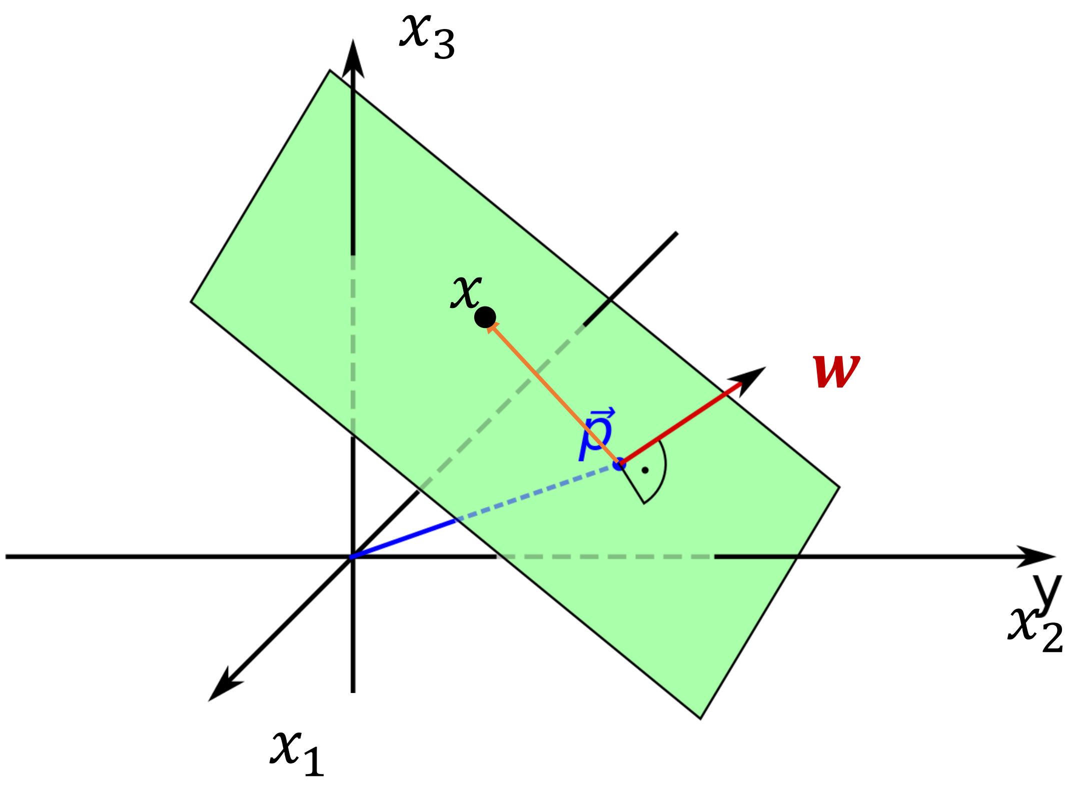


- The generalization of the plane to the m-dimensional space is called the hyperplane and is given by:

$$w_1x_1 + w_2x_2 + \cdots + w_mx_m + w_0 = 0$$

- In a space with m dimensions, a hyperplane is an m – 1 dimensional object (it has no depth)

Defining Hyper-planes

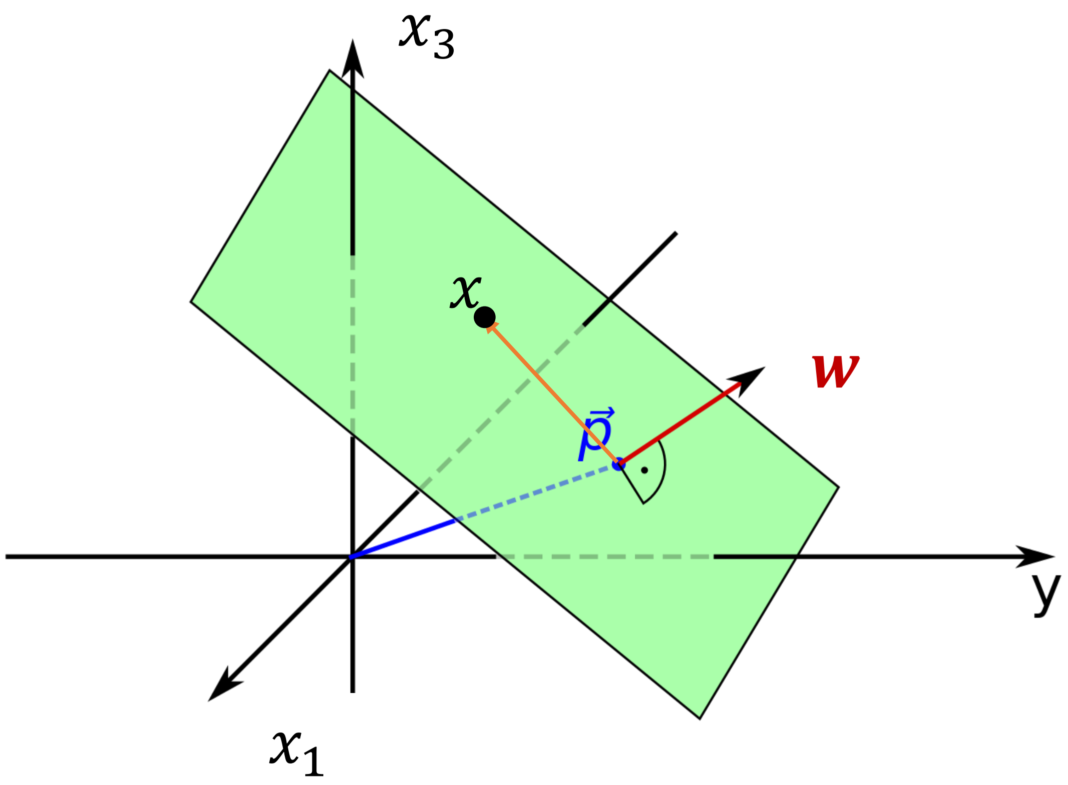


- Planes are defined using "normal form" using a normal vector $\boldsymbol{w}$ and a point $\boldsymbol{p}$ that lies on the plane:

$$\boldsymbol{w} \cdot (\boldsymbol{x} - \boldsymbol{p}) = 0$$

- All points $\boldsymbol{x}$ for which the equation is true lie on the plane
- All other points are not on the plane

Defining Hyperplanes

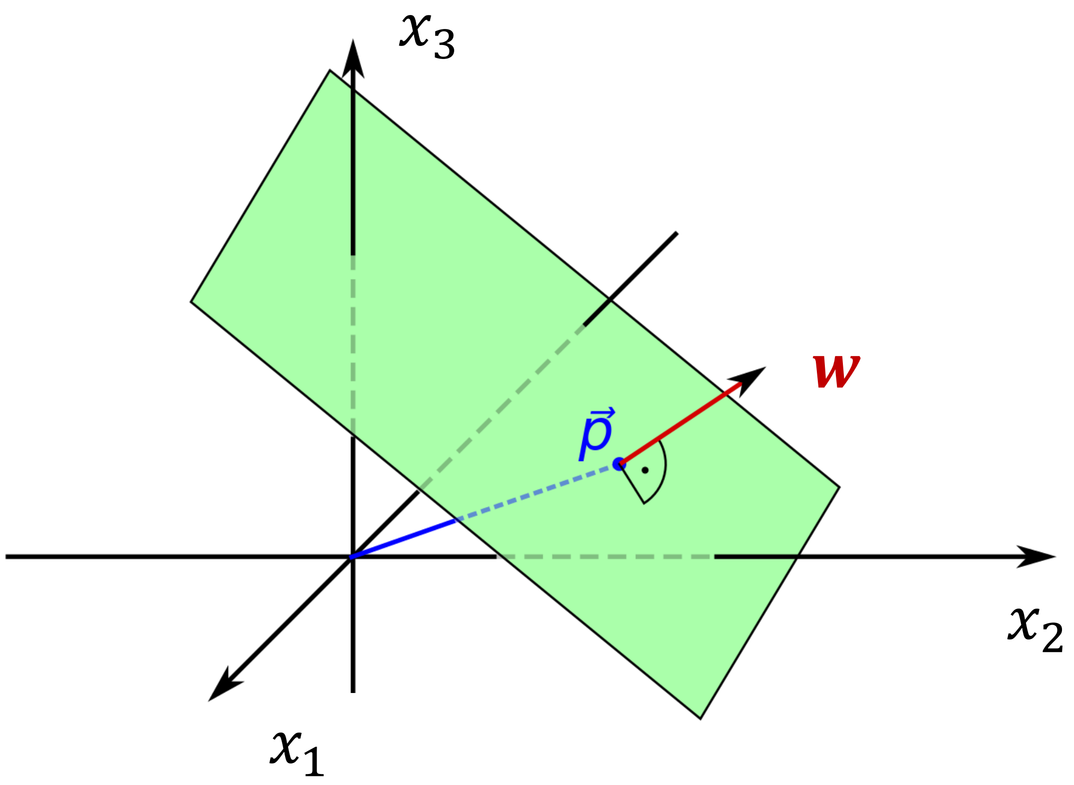


- Note the dot product in the equation:

$$\boldsymbol{w} \cdot (\boldsymbol{x} - \boldsymbol{p}) = 0$$

- The equation measures the angle between the normal vector and a vector between the two points
- If two vectors are orthogonal, then their dot product is 0.

Checking where a point lies relative to a plane



- If Point lies on the plane, then

$$w_1x_1 + w_2x_2 + \cdots + w_mx_m + w_0 = 0$$

- If Point lies on the same side of the plane as the normal vector, then

$$w_1x_1 + w_2x_2 + \cdots + w_mx_m + w_0 > 0$$

- If Point lies on the opposite side of the plane as the normal vector, then

$$w_1x_1 + w_2x_2 + \cdots + w_mx_m + w_0 < 0$$