

Day-II Session-III

Linear Algebra for Data Science



Dr. Sibarama Panigrahi

Department of Computer Science & Engineering
National Institute of Technology Rourkela

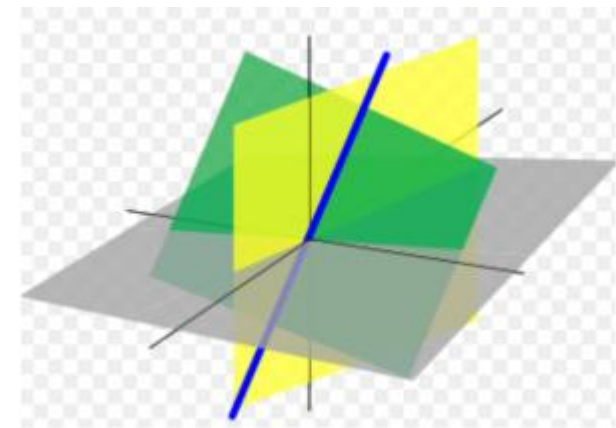
Linear Algebra

- Linear algebra is the branch of mathematics concerning linear equations such as

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b$$

- By default, Vectors are denoted as Column Vectors.
- In vector notation we say $a^T x = b$

$$[a_1, a_2, \dots, a_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = b$$



- Called a linear transformation of x
- Linear algebra is fundamental to geometry, for defining objects such as lines, planes, rotations.

Why Linear Algebra?

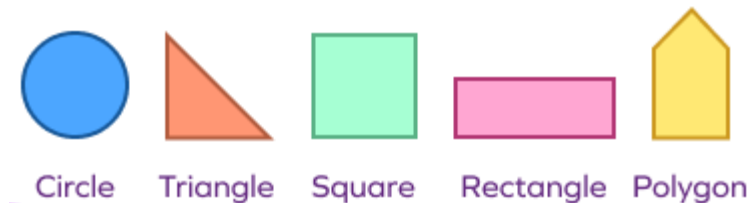
- Linear Algebra provides us the mathematical tool to understand in lower dimensions (2-D/3-D) and generalize for higher dimensions (n-D).

- 0-Dimensional : . (dot)



- 1-Dimensional

- 2-Dimensional

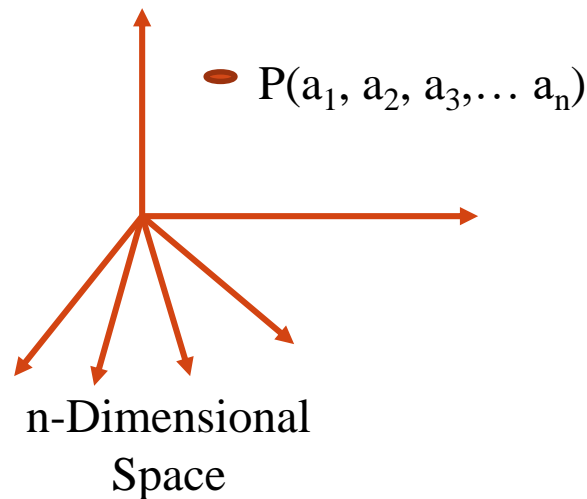
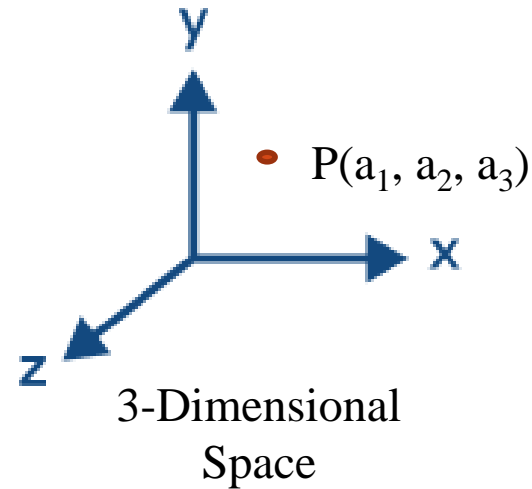
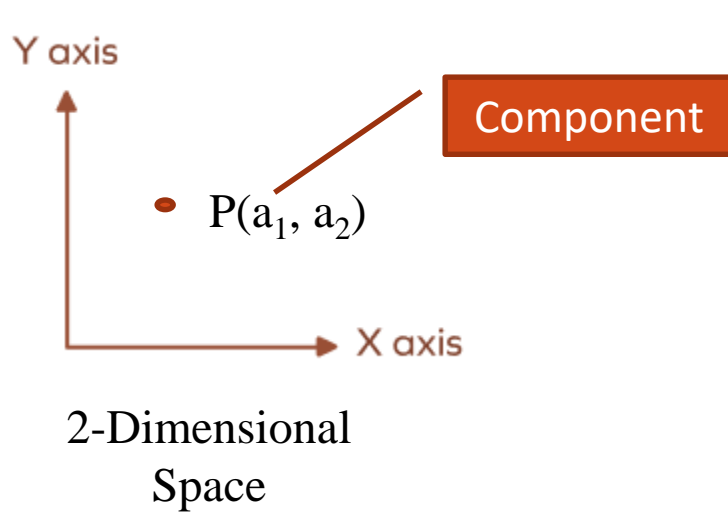


- 3-Dimensional

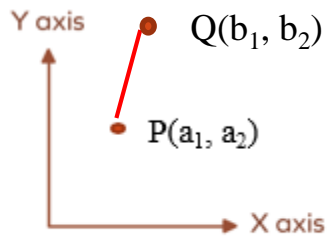


- n-Dimensional: Hypersphere, Hyperplane, Hypercube,...

Point (Vector)

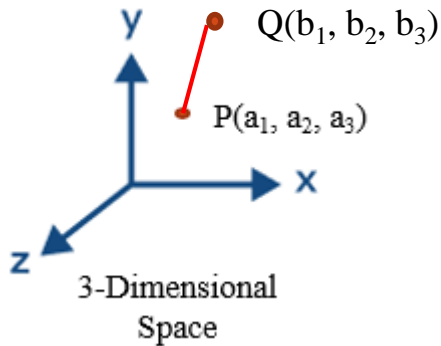


Distance between two Points

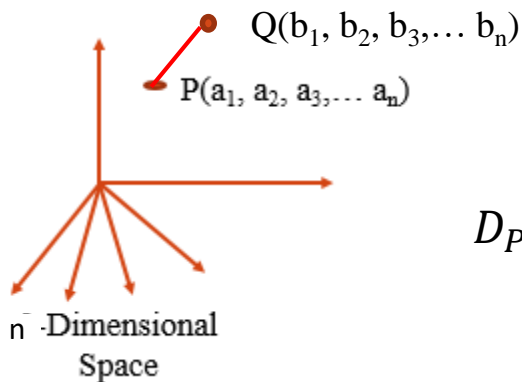


2-Dimensional
Space

$$D_{PQ} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$$

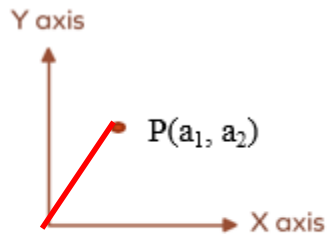


$$D_{PQ} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}$$



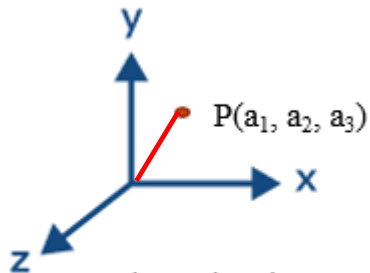
$$D_{PQ} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

Distance of a Point from Origin



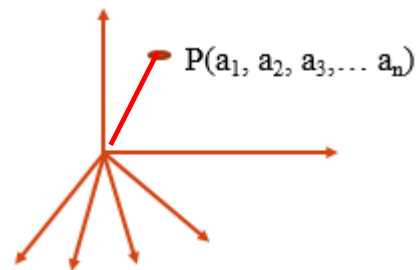
2-Dimensional
Space

$$D = |P| = \sqrt{(a_1 - 0)^2 + (a_2 - 0)^2} = \sqrt{a_1^2 + a_2^2}$$



3-Dimensional
Space

$$D = |P| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$



n-Dimensional
Space

$$D = |P| = \sqrt{a_1^2 + a_2^2 + a_3^2 + \dots + a_n^2} = \sqrt{\sum_{i=1}^n a_i^2}$$

Vector Operations

- $a = [a_1, a_2, \dots, a_n]$
- $b = [b_1, b_2, \dots, b_n]$
- Addition: $a + b = [a_1 + b_1, a_2 + b_2, \dots, a_n + b_n]$
- Subtraction: $a - b = [a_1 - b_1, a_2 - b_2, \dots, a_n - b_n]$
- Multiplication:
 - Dot Product: $a \cdot b = [a_1 b_1 + a_2 b_2 + \dots, a_n b_n]$

$$a \cdot b = [a_1, a_2, \dots, a_n] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = a^T b$$

- Cross Product (Not much used in Data Science)

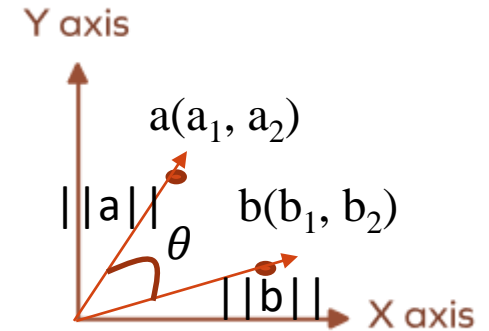
Vector Operations

- Dot Product: (Visualize in 2-D)

- $a \cdot b = \|a\| \|b\| \cos(\theta)$ [Geometry Definition]

- Where $\|a\| = \sqrt{a_1^2 + a_2^2}$ = distance of a from Origin

- $a \cdot b = a_1 b_1 + a_2 b_2$ [Algebra Definition]



- The angle between two vectors $= \theta = \cos^{-1} \left(\frac{a \cdot b}{\|a\| \|b\|} \right)$

$$\theta = \cos^{-1} \left(\frac{a_1 b_1 + a_2 b_2}{\|a\| \|b\|} \right)$$

When $\theta = 90^\circ \rightarrow \cos(90) = 0 \rightarrow a \cdot b = 0$

Vector Operations

- Dot Product: (In n-D)

- $a \cdot b = \|a\| \|b\| \cos(\theta)$

- Where $\|a\| = \sqrt{\sum_{i=1}^n a_i^2}$ = distance of a from Origin

- $a \cdot b = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \sum_{i=1}^n a_i b_i$

- The angle between two vectors = $\theta = \cos^{-1} \left(\frac{a \cdot b}{\|a\| \|b\|} \right)$

$$\theta = \cos^{-1} \left(\frac{\sum_{i=1}^n a_i b_i}{\|a\| \|b\|} \right)$$

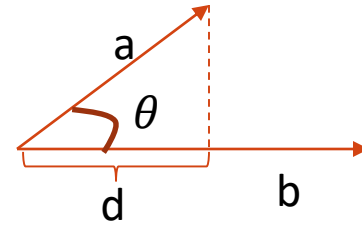
When $\theta = 90^\circ \rightarrow \cos(90) = 0 \rightarrow a \cdot b = 0$

$$a \cdot a = a_1 a_1 + a_2 a_2 + \dots + a_n a_n = \sum_{i=1}^n a_i^2 = \|a\|^2$$

Dot product between two same vectors = (distance from Origin)²

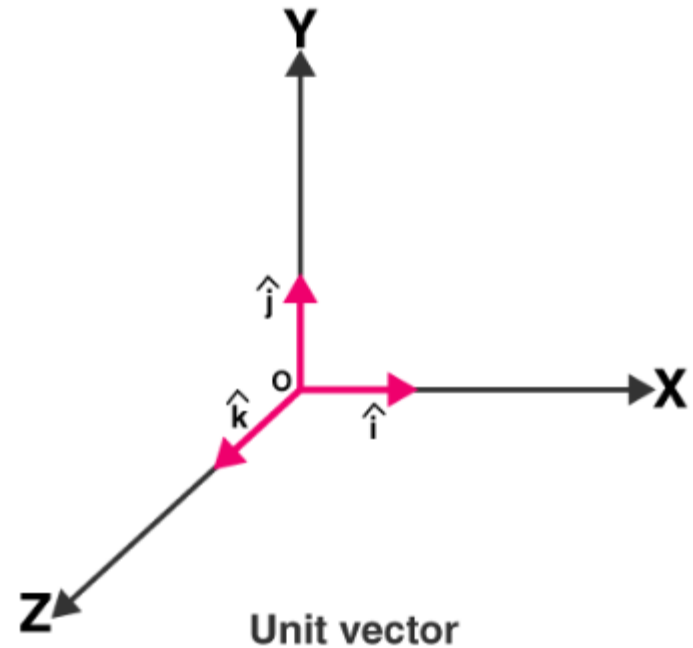
Projection

- $\cos(\theta) = \frac{p}{h} = \frac{d}{\|a\|}$
- *Projection of a on b i.e. $d = \|a\|\cos(\theta)$*
- $d = \frac{a \cdot b}{\|b\|} = \frac{\|a\|\|b\|\cos(\theta)}{\|b\|} = \|a\|\cos(\theta)$

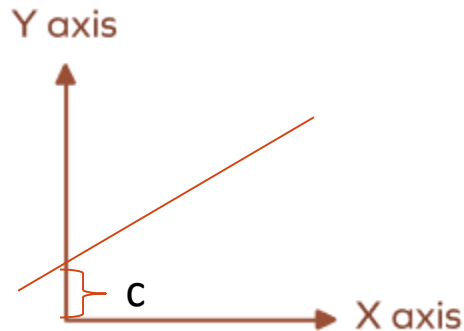


Unit Vector

- A vector is a quantity that has both magnitude, as well as direction.
- A vector that has a magnitude of 1 is a **unit vector**. It is also known as **Direction Vector**.
- Unit vector $\hat{a} = \frac{a}{\|a\|}$



Line (2-D), Plane (3-D) & Hyperplane(n-D)



- 2-D:Line: $y = mx + c$

$$ax + by + c = 0$$

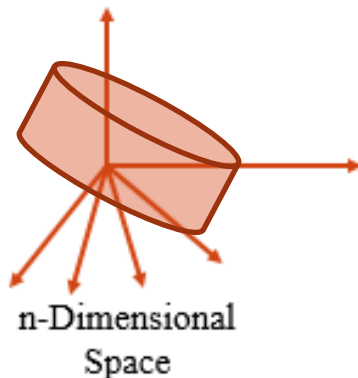
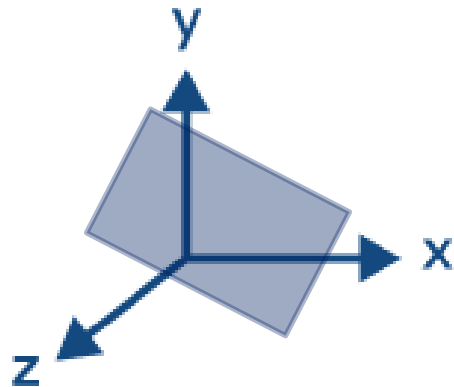
$$w_1x_1 + w_2x_2 + w_0 = 0$$

Note: When $w_0 = 0$, the line passes through the origin.

- 3-D: Plane:

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

0



- n-D:Hyperplane

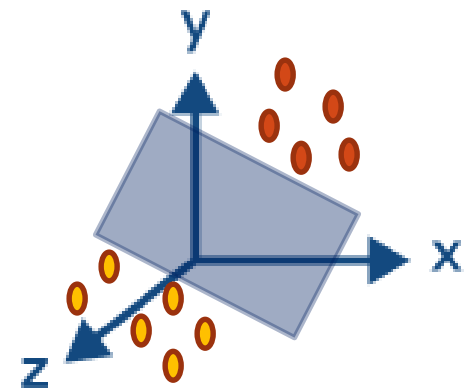
$$w_1x_1 + \dots + w_nx_n + w_0 = 0$$

$$w \cdot x + w_0 = 0$$

When $w_0 = 0$, the hyperplane passes through origin.

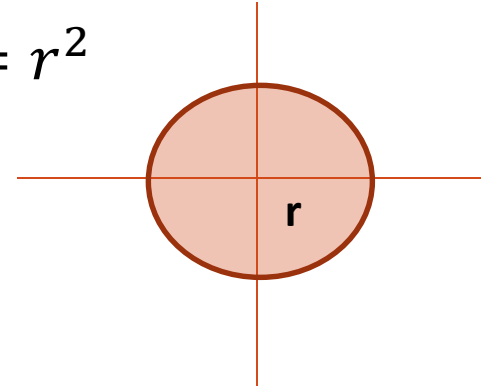
Line (2-D), Plane (3-D) & Hyperplane(n-D)

- A plane breaks a plane into two half-spaces one above the plane and one below the plane
- If the dot product of plane (w) and point (p) is positive then the point p is lying in the same half plane as that of w otherwise it is lying in the other half plane (i.e. opposite direction to w)



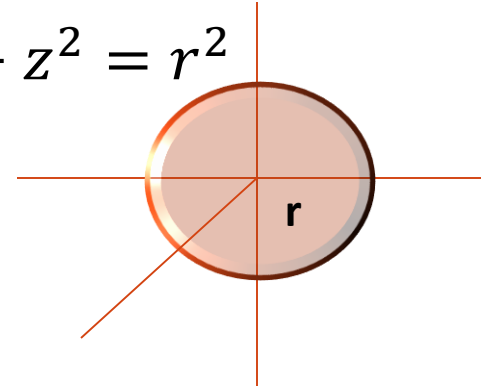
Circle (2-D), Sphere (3-D) & Hypersphere(n-D)

- 2-D (Circle)
- Equation of a circle centred at origin: $x^2 + y^2 = r^2$
- A point $p(x_1, x_2)$ lies
 - Inside the circle if $x_1^2 + x_2^2 < r^2$
 - Outside the circle if $x_1^2 + x_2^2 > r^2$
 - On the circle if $x_1^2 + x_2^2 = r^2$



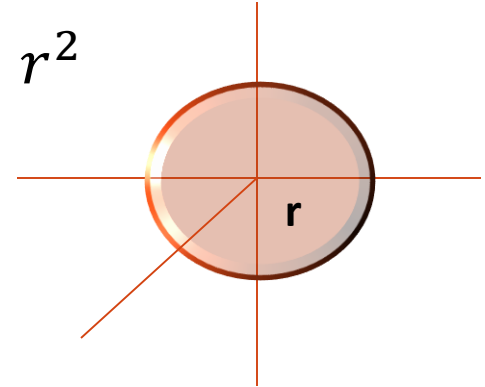
Circle (2-D), Sphere (3-D) & Hypersphere(n-D)

- 3-D (Sphere)
- Equation of a circle centred at origin: $x^2 + y^2 + z^2 = r^2$
- A point $p(x_1, x_2, x_3)$ lies
 - Inside the circle if $x_1^2 + x_2^2 + x_3^2 < r^2$
 - Outside the circle if $x_1^2 + x_2^2 + x_3^2 > r^2$
 - On the circle if $x_1^2 + x_2^2 + x_3^2 = r^2$



Circle (2-D), Sphere (3-D) & Hypersphere(n-D)

- n-D (Hypersphere)
- Equation of a circle centred at origin: $\sum_{i=1}^n x_i^2 = r^2$
- A point $p(x_1, x_2, x_3 \dots x_n)$ lies
 - Inside the circle if $\sum_{i=1}^n x_i^2 < r^2$
 - Outside the circle if $\sum_{i=1}^n x_i^2 > r^2$
 - On the circle if $\sum_{i=1}^n x_i^2 = r^2$



Ellipse, Ellipsoid, Hyperellipsoid

■ *Ellipse*

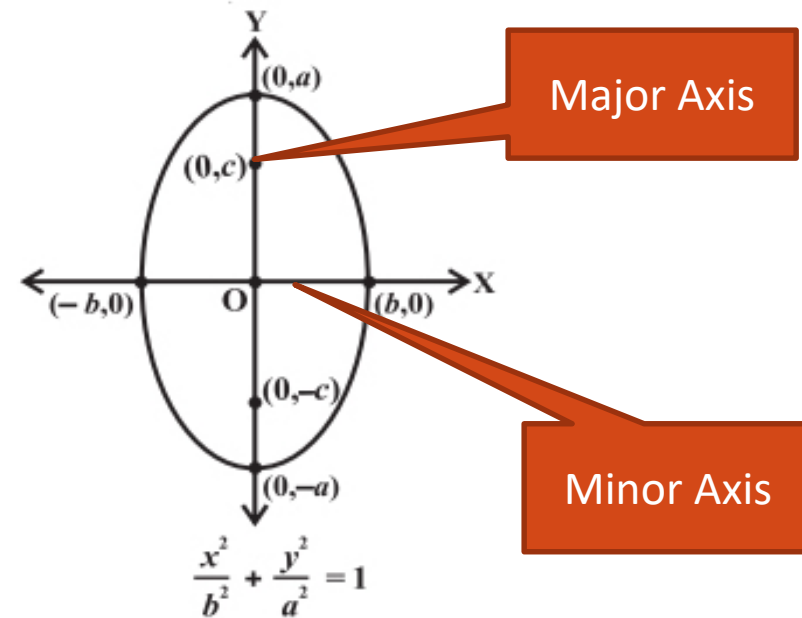
- $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ point $p(x,y)$ lies on the ellipse
- $\frac{x^2}{a^2} + \frac{y^2}{b^2} < 1$ point $p(x,y)$ lies inside ellipse
- $\frac{x^2}{a^2} + \frac{y^2}{b^2} > 1$ point $p(x,y)$ lies outside the ellipse

■ *Ellipsoid(3-D)*

- $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$

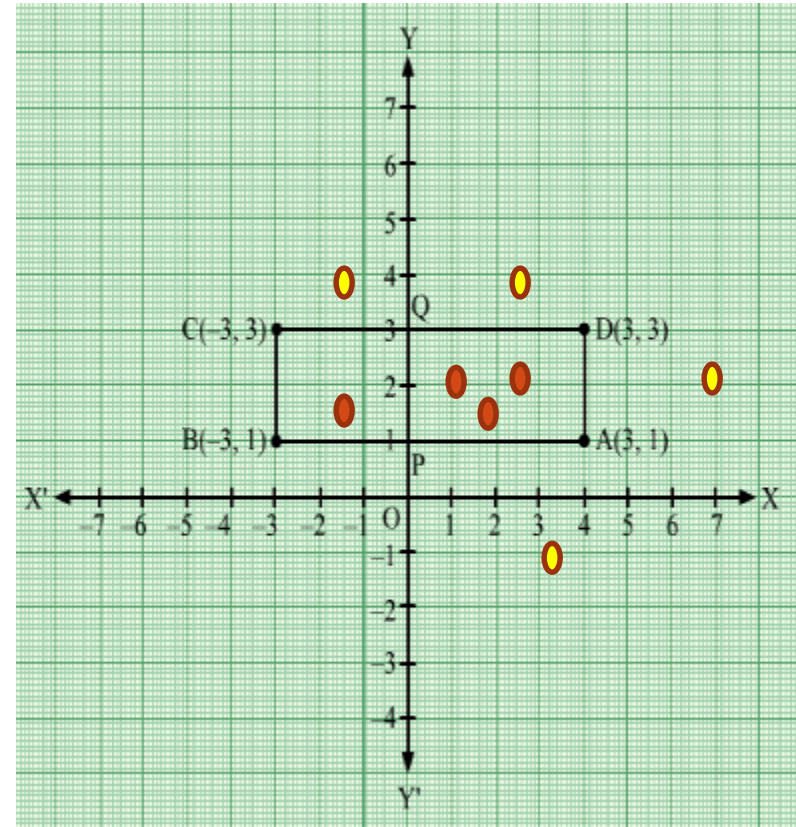
■ *Hyperellipsoid(n-D)*

- *Similarly*



Rectangle, Hyperrectangle, Square, Cube, Hypercube

- If $x \geq -3$ & $x \leq 3$
If $y \geq 1$ & $y \leq 3$
then point $p(x,y)$
lies within the rectangle

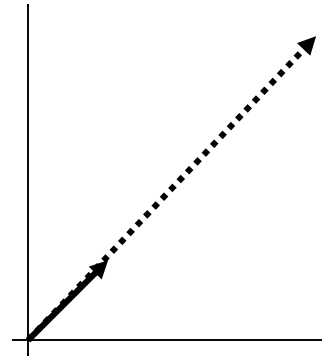


- **Eigenvectors & eigenvalues**

Introduction to Linear Algebra
by Mark Goldman, and Mily Mackevicius

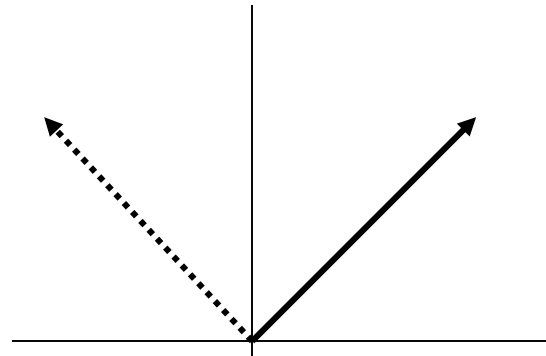
Matrices as linear transformations

$$\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$



(stretching)

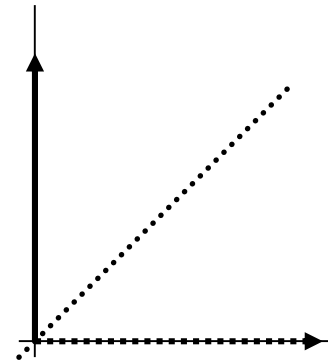
$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$



(rotation)

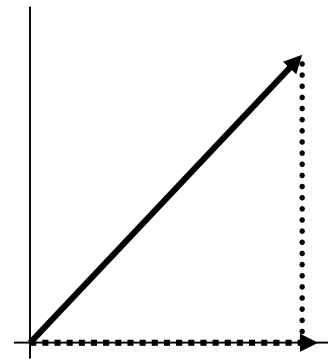
Matrices as linear transformations

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$



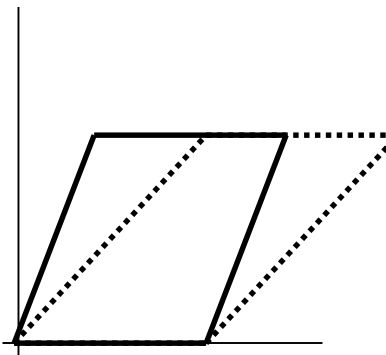
(reflection)

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$



(projection)

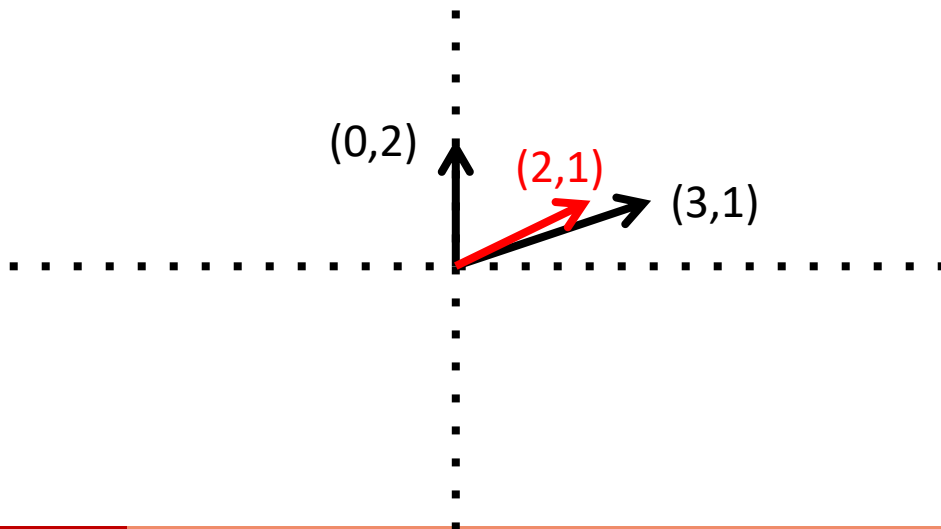
$$\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + cy \\ y \end{pmatrix}$$



(shearing)

What do matrices do to vectors?

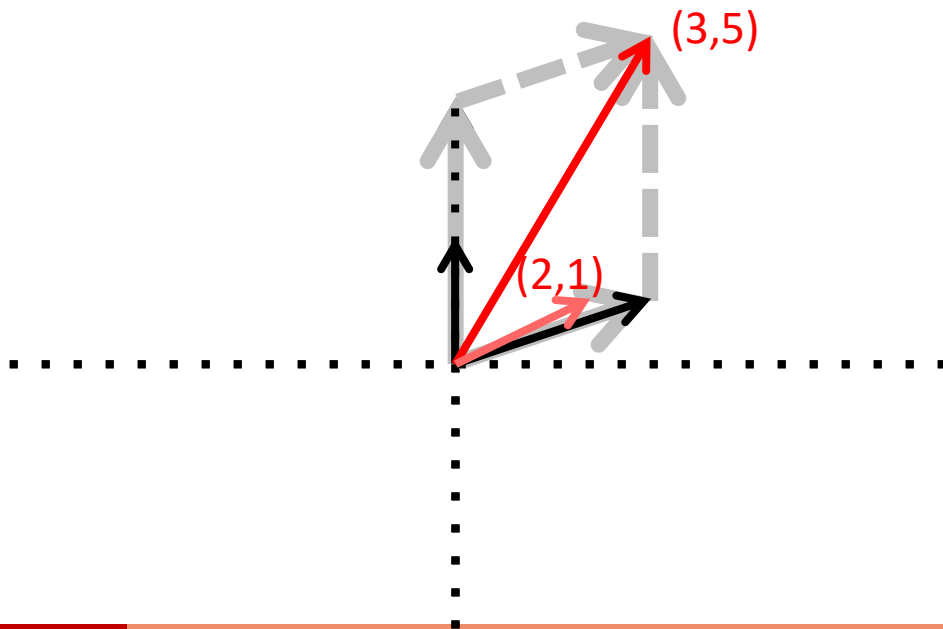
$$\begin{matrix} \overleftrightarrow{M} \\ \swarrow \\ \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \end{matrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} =$$



Recall

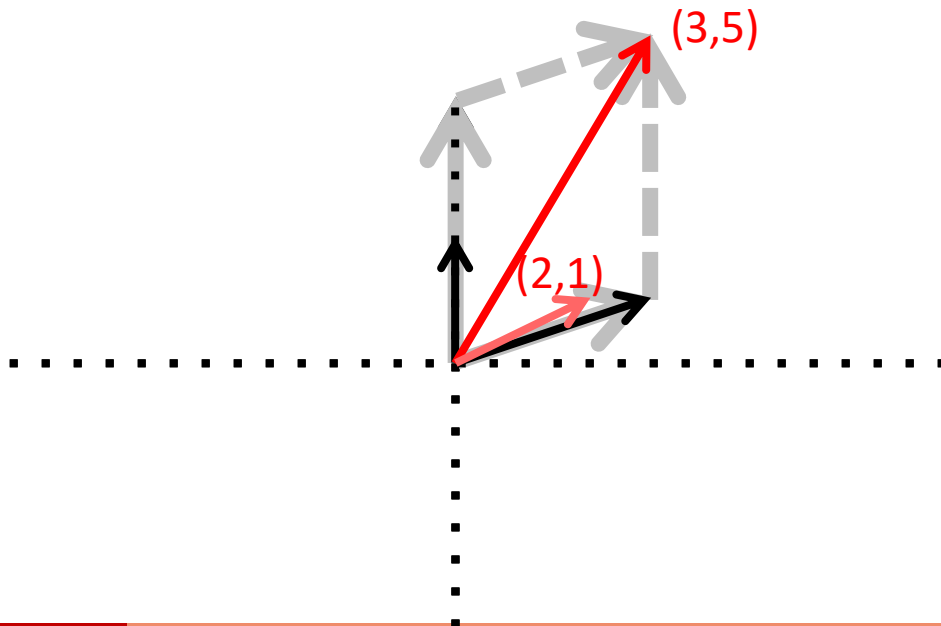
\overleftarrow{M}

$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 1 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$



What do matrices do to vectors?

$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 1 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$



- The new vector is:
 - 1) **rotated**
 - 2) **scaled**

Are there any special vectors that **only** get scaled?

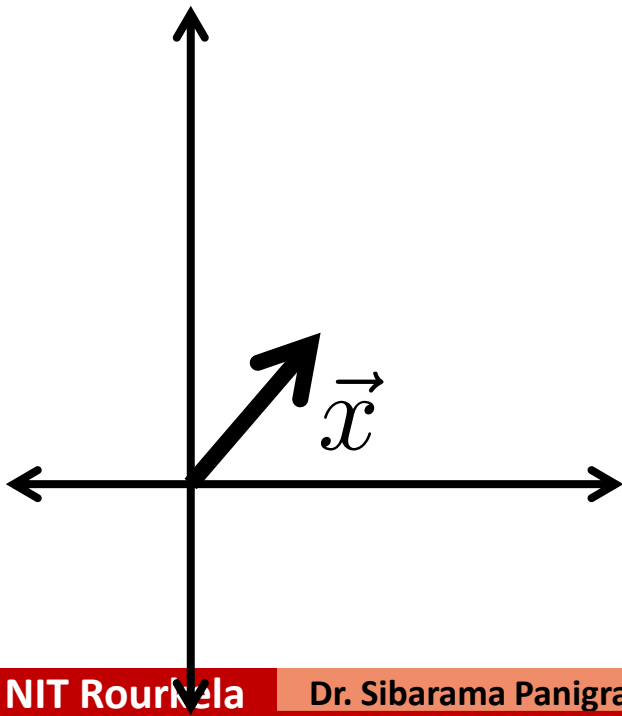
$$\vec{M} \rightarrow \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix}$$

Are there any special vectors that **only** get scaled?



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

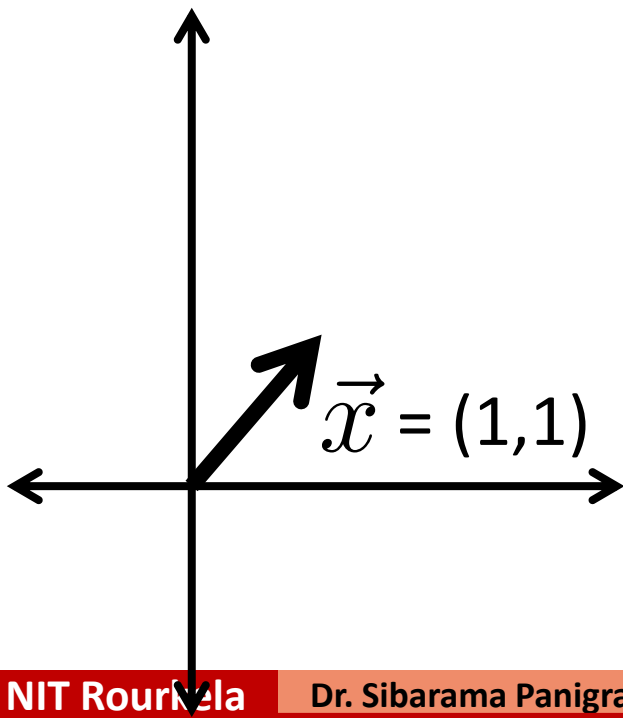
Try (1,1)



Are there any special vectors that **only** get scaled?



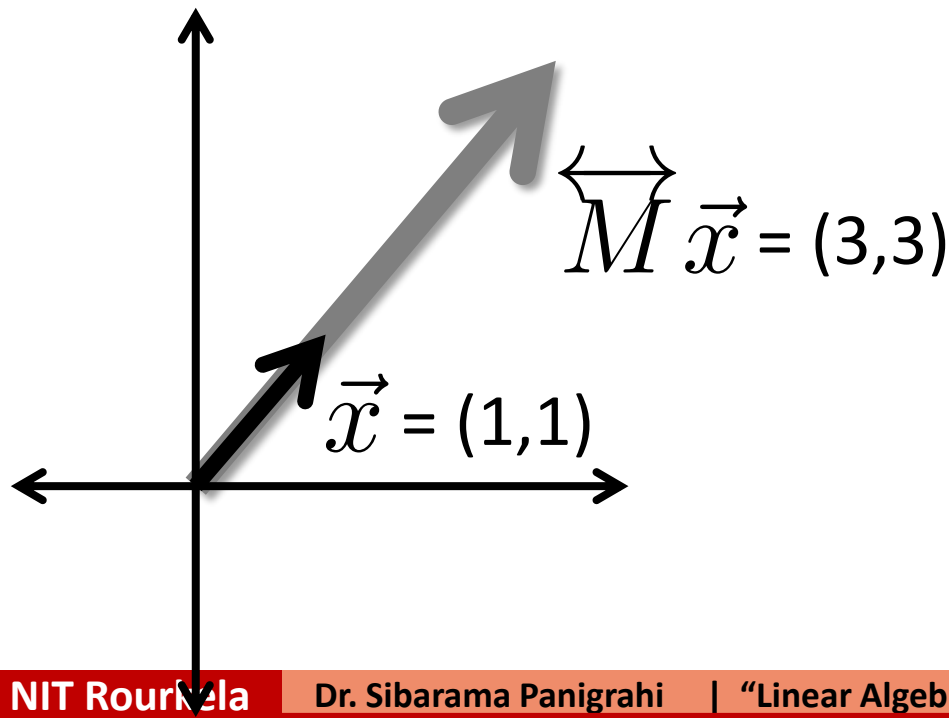
$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$



Are there any special vectors that **only** get scaled?



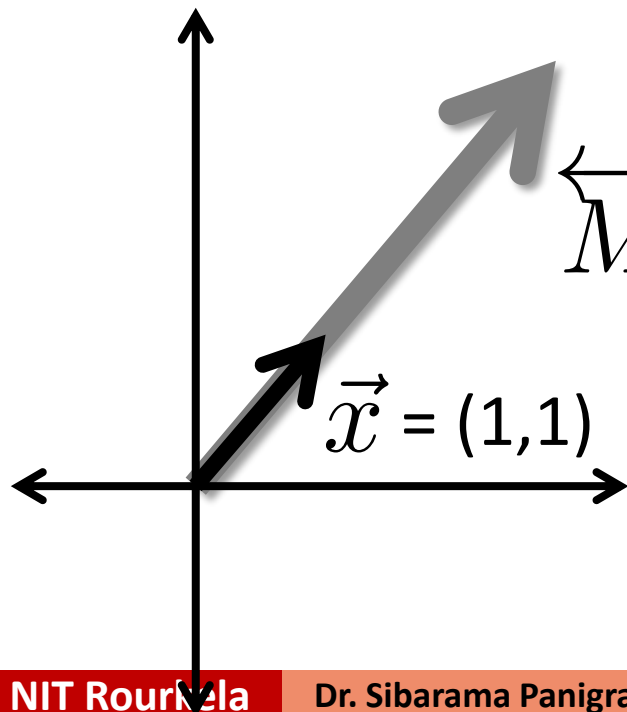
$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



Are there any special vectors that **only** get scaled?



$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



$$M \vec{x} = (3, 3)$$

- For this special vector, multiplying by M is like multiplying by a scalar.
- $(1, 1)$ is called an **eigenvector** of M
- 3 (the scaling factor) is called the **eigenvalue** associated with this eigenvector

Are there any other eigenvectors?

- Yes! The easiest way to find is with Python's eig()

```
import numpy as np
from numpy.linalg import eig
```

```
a = np.array([[0, 2],
              [2, 3]])
w,v=eig(a)
print('E-value:', w)
print('E-vector', v)
```

```
E-value: [-1.  4.]
E-vector [[-0.89442719 -0.4472136 ]
          [ 0.4472136  -0.89442719]]
```

- **Vector space**

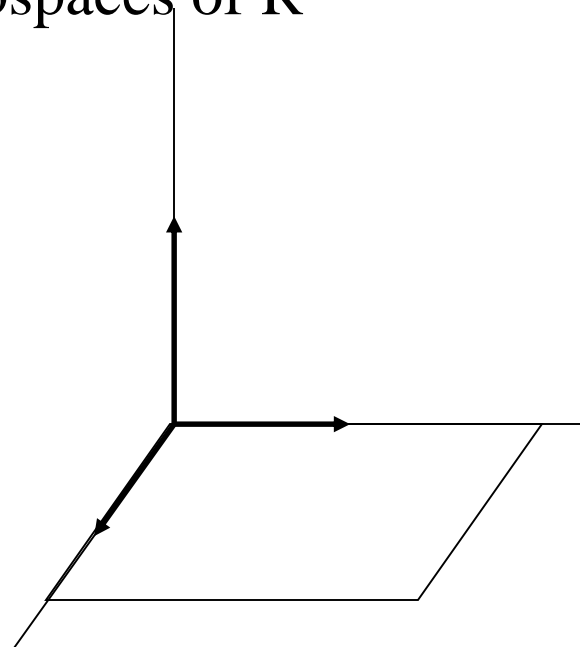
- 10-725 – Optimization 1/16/08 Recitation
- Joseph Bradley

Vector spaces

- Formally, a *vector space* is a set of vectors which is closed under addition and multiplication by real numbers.
- A *subspace* is a subset of a vector space which is a vector space itself, e.g. the plane $z=0$ is a subspace of \mathbb{R}^3 (It is essentially \mathbb{R}^2 .).
- We'll be looking at \mathbb{R}^n and subspaces of \mathbb{R}^n

Our notion of planes in \mathbb{R}^3 may be extended to *hyperplanes* in \mathbb{R}^n (of dimension $n-1$)

Note: subspaces must include the origin (zero vector).

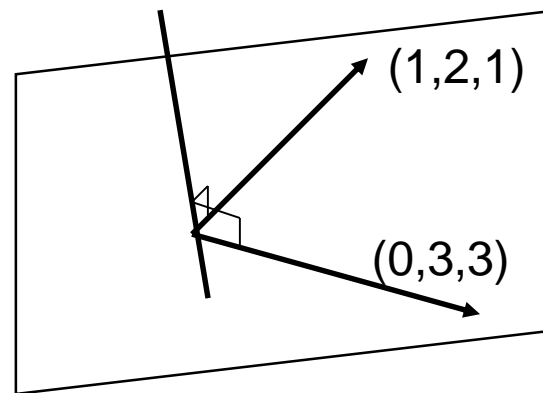


Linear system & subspaces

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$u \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + v \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

- Linear systems define certain subspaces
- $Ax = b$ is solvable iff b may be written** as a linear combination of the columns of A .
- The set of possible vectors b forms a subspace called the column space of A .



Linear system & subspaces

The set of solutions to $Ax = 0$ forms a subspace called the *null space* of A .

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

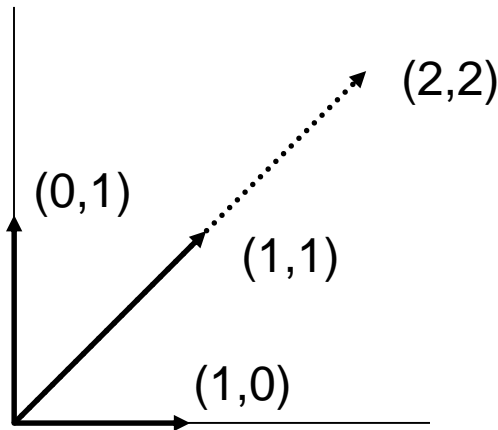
→ Null space: $\{(0,0)\}$

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 3 & 5 \\ 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

→ Null space: $\{(c, c, -c)\}$

Linear independence and basis

- Vectors v_1, \dots, v_k are linearly independent if $c_1 v_1 + \dots + c_k v_k = 0$ implies $c_1 = \dots = c_k = 0$
i.e. the nullspace is the origin



$$\begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

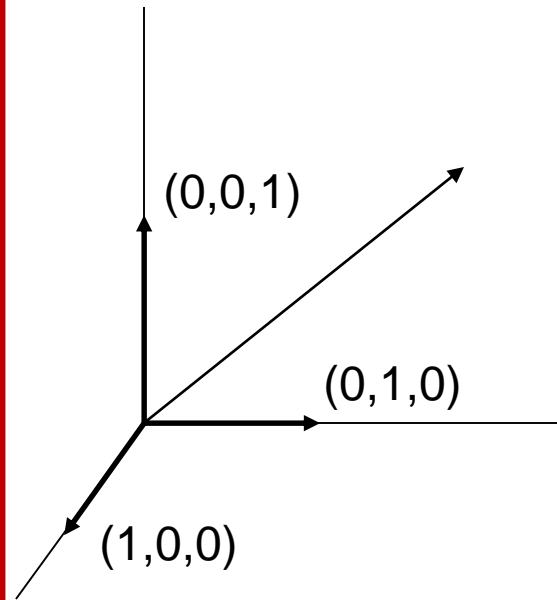
$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Recall nullspace contained only $(u,v)=(0,0)$.
i.e. the columns are linearly independent.

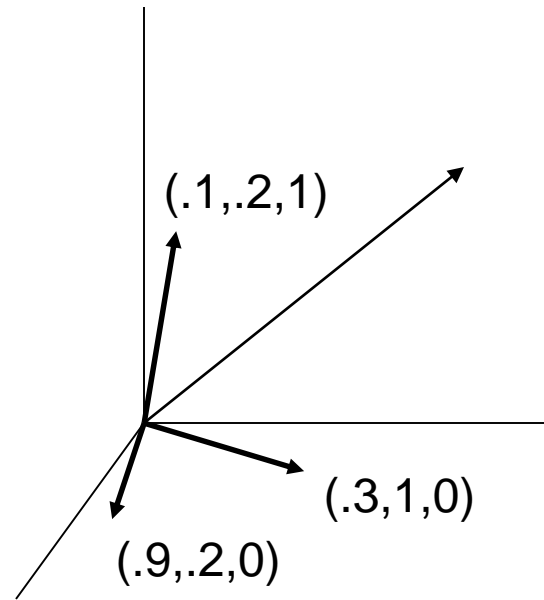
Linear independence and basis

- If all vectors in a vector space may be expressed as linear combinations of v_1, \dots, v_k , then v_1, \dots, v_k *span* the space.

$$\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

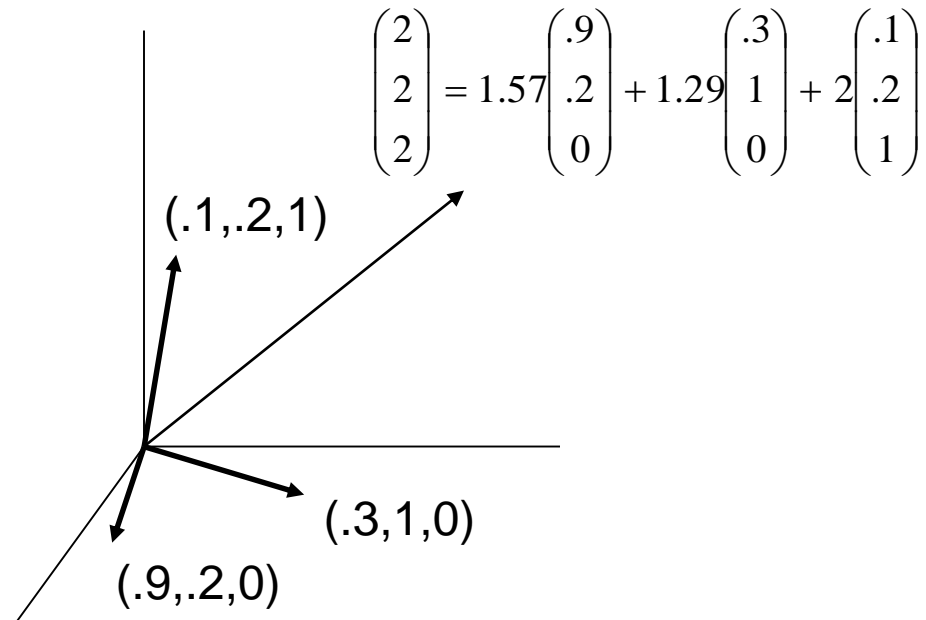
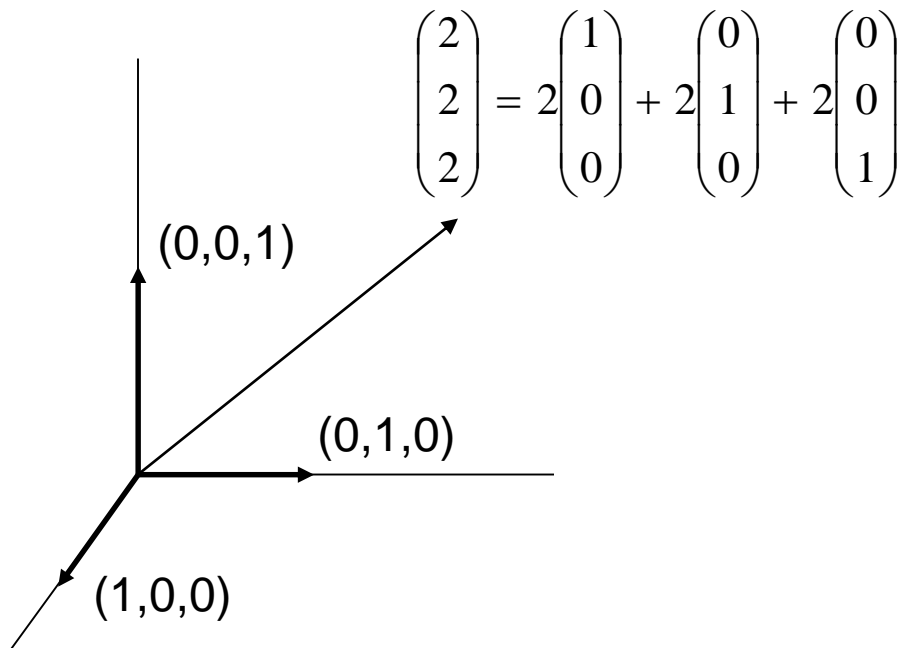


$$\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 1.57 \begin{pmatrix} .9 \\ .2 \\ 0 \end{pmatrix} + 1.29 \begin{pmatrix} .3 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} .1 \\ .2 \\ 1 \end{pmatrix}$$



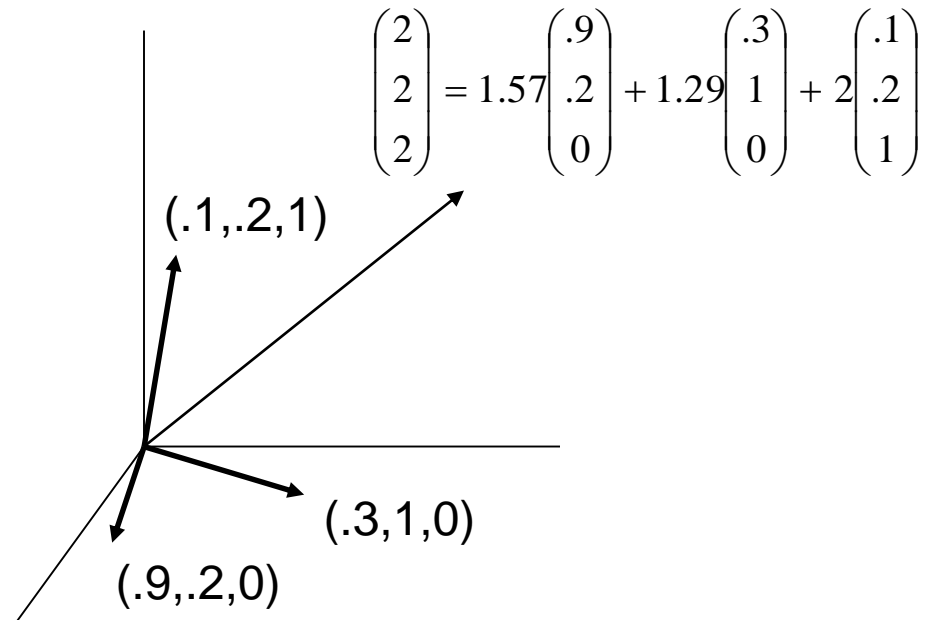
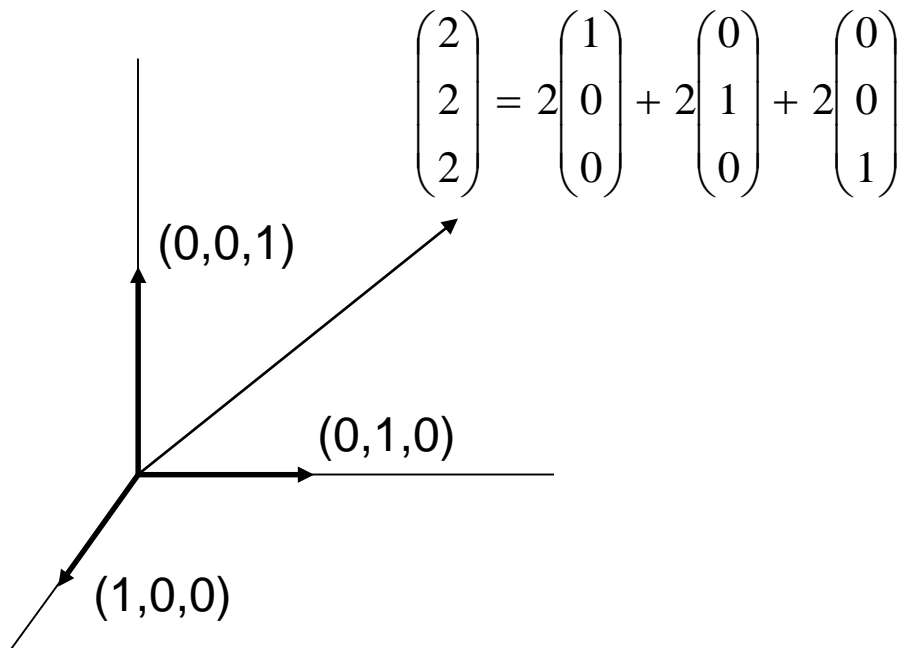
Linear independence and basis

- A *basis* is a set of linearly independent vectors which span the space.
- The *dimension* of a space is the # of “degrees of freedom” of the space; it is the number of vectors in any basis for the space.
- A basis is a maximal set of linearly independent vectors and a minimal set of spanning vectors.



Linear independence and basis

- Two vectors are *orthogonal* if their dot product is 0.
- An *orthogonal basis* consists of orthogonal vectors.
- An *orthonormal basis* consists of orthogonal vectors of unit length.



About subspaces

- The *rank* of A is the dimension of the column space of A .
- It also equals the dimension of the *row space* of A (the subspace of vectors which may be written as linear combinations of the rows of A).

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

$$(1,3) = (2,3) - (1,0)$$

Only 2 linearly independent rows, so rank = 2.

About subspaces

Fundamental Theorem of Linear Algebra:

If A is $m \times n$ with rank r ,

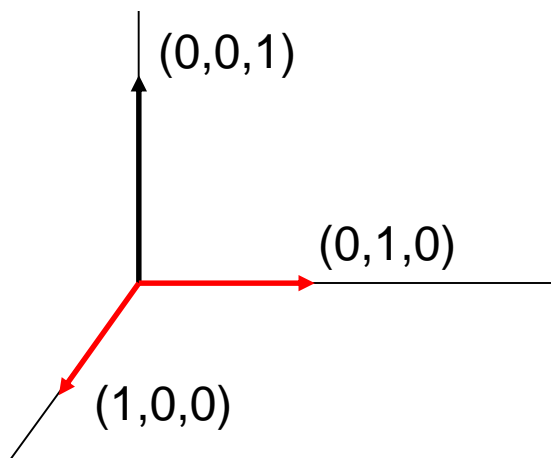
Column space(A) has dimension r

Nullspace(A) has dimension $n-r$ ($=$ *nullity* of A)

Row space(A) = Column space(A^T) has dimension r

Left nullspace(A) = Nullspace(A^T) has dimension $m - r$

Rank-Nullity Theorem: rank + nullity = n



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

$$m = 3$$

$$n = 2$$

$$r = 2$$

Non-square matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 3 \end{pmatrix} \quad \begin{array}{l} m = 3 \\ n = 2 \\ r = 2 \end{array} \quad \begin{array}{l} \text{System } Ax=b \text{ may not} \\ \text{have a solution (x has} \\ \text{2 variables but 3} \\ \text{constraints).} \end{array} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix} \quad \begin{array}{l} m = 2 \\ n = 3 \\ r = 2 \end{array} \quad \begin{array}{l} \text{System } Ax=b \text{ is} \\ \text{underdetermined (x} \\ \text{has 3 variables and 2} \\ \text{constraints).} \end{array} \quad \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

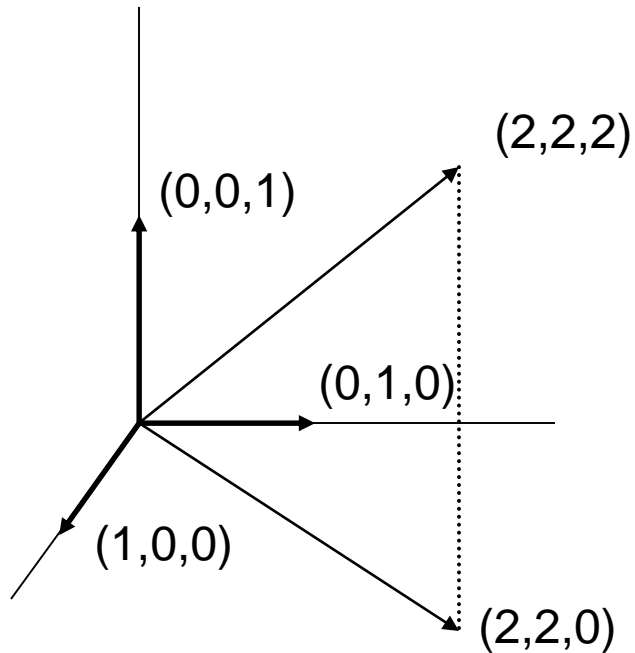
Basis transformations

- Before talking about basis transformations, we need to recall matrix inversion and projections.

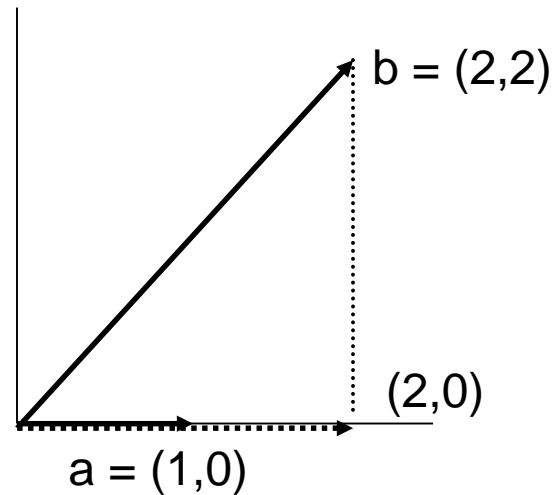
Matrix inversion

- To solve $Ax=b$, we can write a closed-form solution if we can find a matrix A^{-1}
s.t. $AA^{-1}=A^{-1}A=I$ (identity matrix)
- Then $Ax=b$ iff $x=A^{-1}b$:
$$x = Ix = A^{-1}Ax = A^{-1}b$$
- A is *non-singular* iff A^{-1} exists iff $Ax=b$ has a unique solution.
- Note: If A^{-1}, B^{-1} exist, then $(AB)^{-1} = B^{-1}A^{-1}$,
and $(A^T)^{-1} = (A^{-1})^T$

Projections



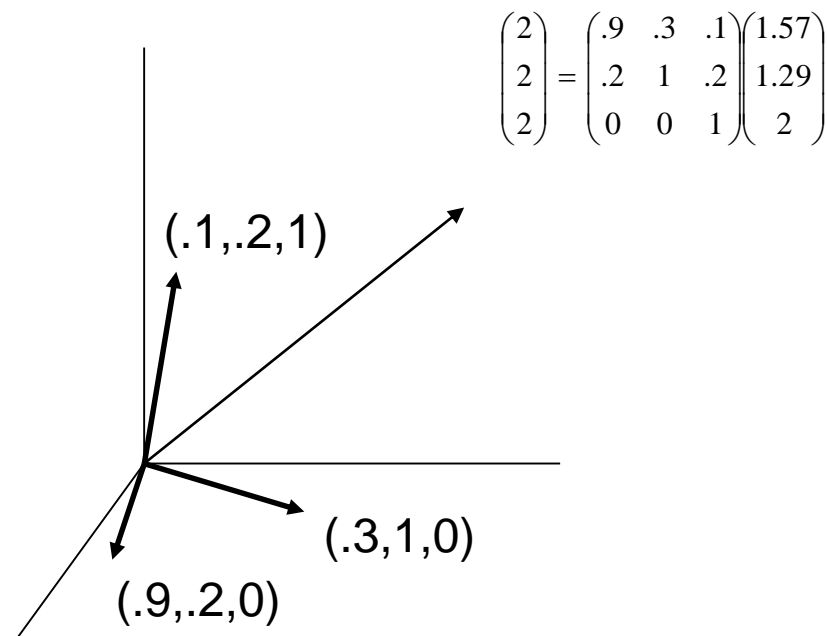
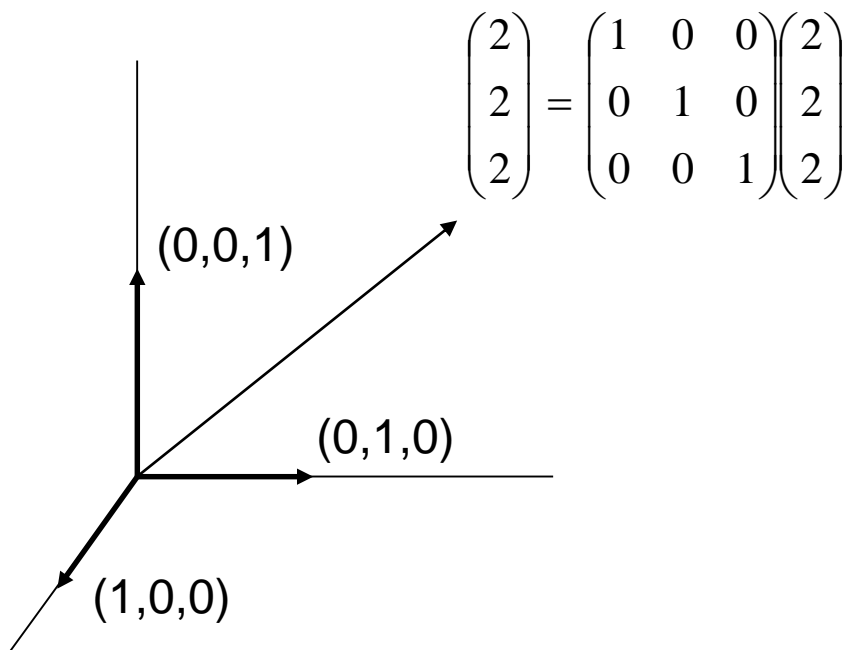
$$\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$



$$c = \frac{a^T b}{a^T a} a = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

Basis transformations

We may write $v=(2,2,2)$ in terms of an alternate basis:



Components of $(1.57, 1.29, 2)$ are projections of v onto new basis vectors, normalized so new v still has same length.

Determinants

- If $\det(A) = 0$, then A is singular.
- If $\det(A) \neq 0$, then A is invertible.
- To compute:
 - Simple example:
 - Python: `det(A)`

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

Determinants

- m -by- n matrix A is *rank-deficient* if it has rank $r < m$ ($\leq n$)
- Thm: $\text{rank}(A) < r$ iff
$$\det(A) = 0 \text{ for all } t\text{-by-}t \text{ submatrices,}$$
$$r \leq t \leq m$$

- **Thank You**