Applied Machine Learning

Dr. Harikrishnan N B
Computer Science and Information Systems

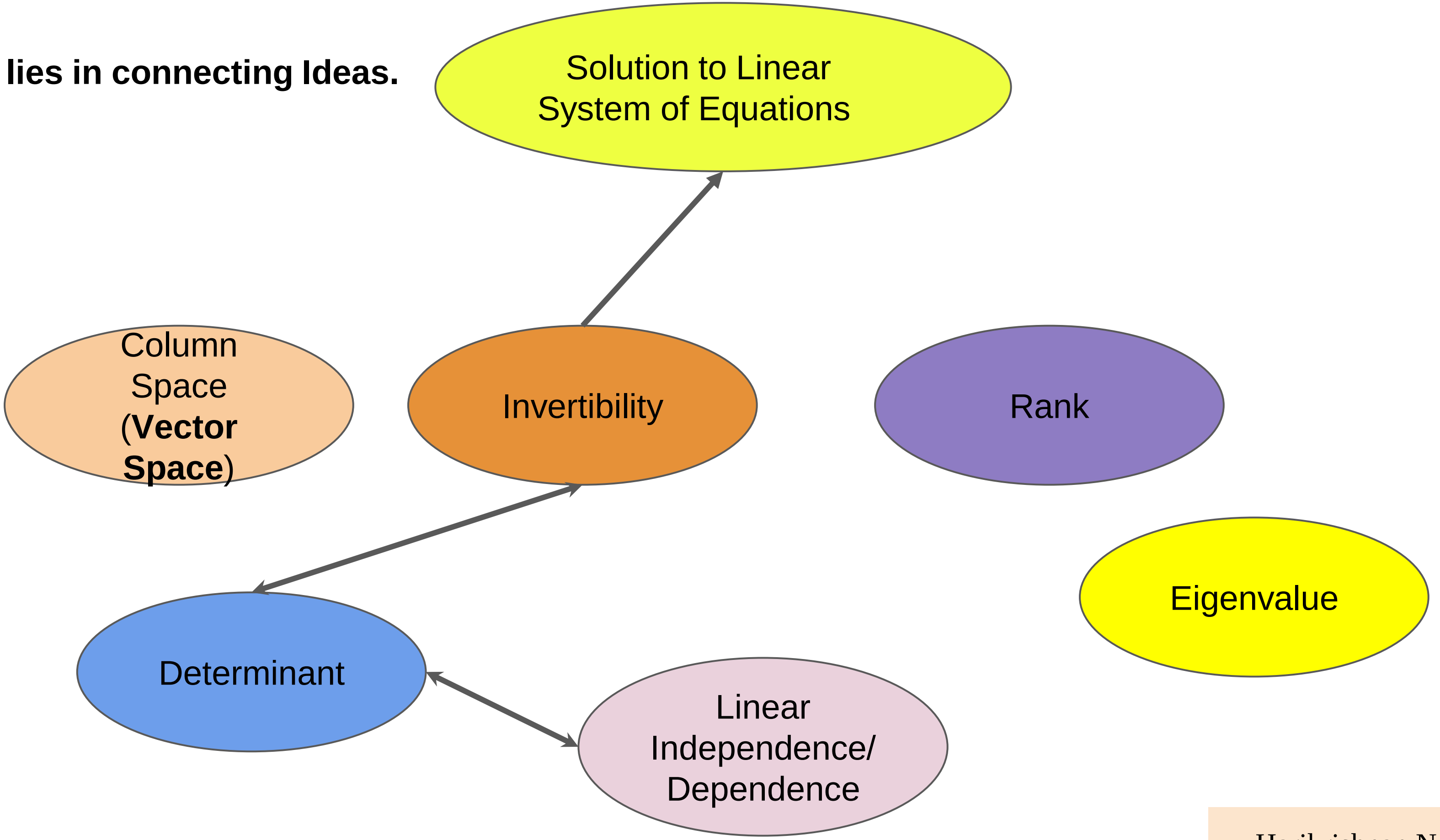
BITS Pilani
Pilani Campus

SE ZG568 / SS ZG568,
Applied Machine Learning
Lecture No. 4 [09- Feb-2025]

Recap

Basics of Linear Algebra,
Row Picture, Col Picture, Algebraic Way
Solution to System of Linear Equations
Inverse of a Matrix
Linear Regression, PCA

Beauty lies in connecting Ideas.



Orthogonal and Orthonormal Matrix

Two vectors are orthogonal when their dot product is zero

$\vec{a} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \vec{b} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

Orthogonal vectors	Orthonormal vectors

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

L2 - norm = $\sqrt{2}$

$\vec{a} \cdot \vec{b} = 0$

$\Rightarrow \|a\|_2 \|b\|_2 \cos\theta = 0$

$\cos\theta = 0$ ✓

$\theta = 90^\circ$ ✓

$\boxed{\|a\|_2, \|b\|_2 \neq 0}$

$a = \begin{pmatrix} a_x \\ a_y \end{pmatrix}$ $b = \begin{pmatrix} b_x \\ b_y \end{pmatrix}$

L2 norm of $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow \sqrt{1^2 + 1^2} = \sqrt{2}$

90°

$\vec{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ = L2 norm of $\vec{a} = \sqrt{1^2 + 2^2} = \sqrt{5}$

$\begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$ L2 norm of $\sqrt{\left(\frac{1}{\sqrt{5}}\right)^2 + \left(\frac{2}{\sqrt{5}}\right)^2} = \sqrt{5/5} = 1$

L2 norm of $\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

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

$\sqrt{\left(1/\sqrt{2}\right)^2 + \left(-1/\sqrt{2}\right)^2}$ $= \sqrt{\frac{1}{2} + \frac{1}{2}} = 1$

L2 - norm = 1

$\boxed{\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \vec{b} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}}$

$a^T b = a \cdot b$

$\Rightarrow \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = 1 \times (-2) + 2 \times 1 = 0$

Orthonormal Matrix

$O(N^3)$

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\quad A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\begin{array}{cc} \text{col}_1 & \text{col}_2 \end{array}$$
$A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

$\text{col}_1 \cdot \text{col}_2 = 0$

$\|\text{col}_1\|_2 = 1, \quad \|\text{col}_2\|_2 = 1$

$|A| = (ad - bc)$

$\frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \times -\frac{1}{\sqrt{2}}$

$\frac{1}{2} + \frac{1}{2} = 1$

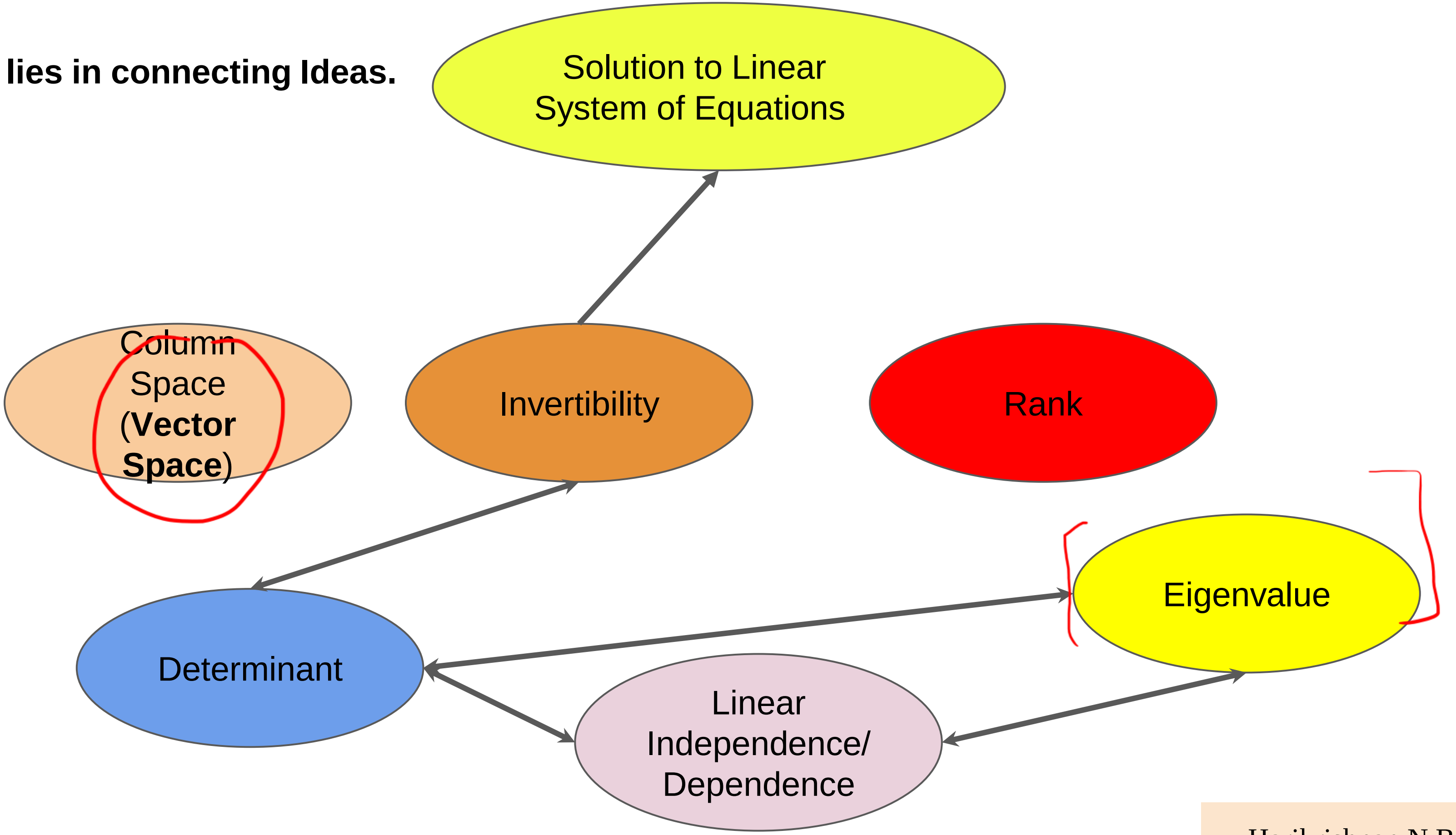
$$\begin{array}{cccc} \text{col}_1 & \text{col}_2 & \cdots & \text{col}_n \end{array}$$
$\begin{bmatrix} | & | & \cdots & | \\ | & | & & | \end{bmatrix}$

for $i \neq j$
$\text{col}_i \cdot \text{col}_j = 0$
for all i from 1 to N
$\|\text{col}_i\|_2 = 1$

$$\begin{array}{cc} \text{col}_1 & \text{col}_2 \end{array}$$
$A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ $\quad A^{-1} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

$A^{-1} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

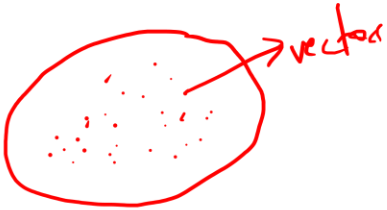
Beauty lies in connecting Ideas.



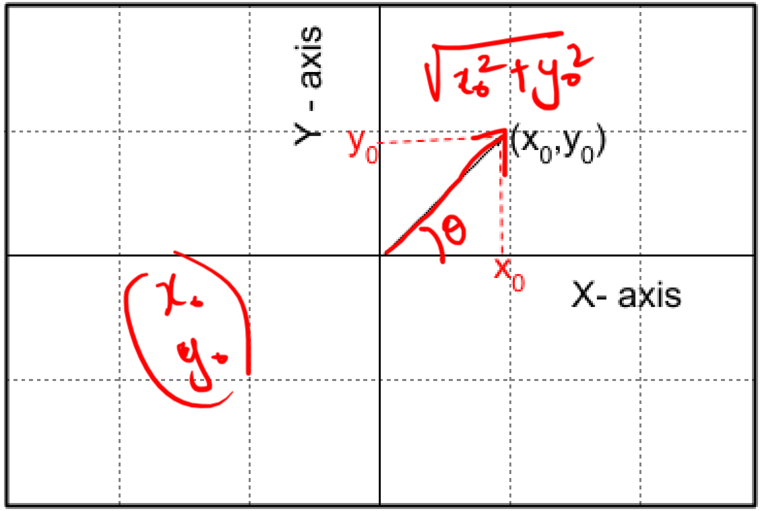
Third Iteration

Harikrishnan N B

Vectors - Different Understanding



Physicists



Computer Scientist

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

Mathematicians

Vector space is a **collection of objects**(it can be anything) called vectors which satisfies mainly two important properties:

1. **closed under vector addition** ✓
2. **closed under scalar multiplication.** ✓

Vector Space - Coffee Space

Coffee space = { coffee$_1$, coffee$_2$ . . coffee$_1$ coffee$_N$ . . . }

$\mathbb{R} = \{ \ldots 1.2, \ldots 2.5 \}$

$1.2 + 2.5 = 3.7$

Coffee Space - In Coffee space we have different kinds of coffee with varying strength. Now we will understand the vector space properties with this metaphor.

$x \in \mathbb{R}^2$, y

$a \in \mathbb{R}^2$

$a = \begin{pmatrix} a_x \\ a_y \end{pmatrix}$

$\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \text{ such that } x \in \mathbb{R}, y \in \mathbb{R} \right\}$

Closed under Vector Addition

$\vec{b} = -\vec{a}$

$b \in \mathbb{R}^2$

$\vec{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ $\vec{b} \begin{pmatrix} -2 \\ -3 \end{pmatrix}$ $b = \begin{pmatrix} b_x \\ b_y \end{pmatrix}$

$a + b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\mathbb{R}^2$ and remove the vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Closed under Scalar Multiplication

$\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \text{ such } x, y \in \mathbb{R} \right\}$

$c \begin{pmatrix} x \\ y \end{pmatrix}$ where $c \in \mathbb{R}$

$\mathbb{R}^2$ is a valid space

Adding two coffee's will give you another coffee which is in the coffee space

Scaling a coffee will give a coffee which is in the coffee space

$\mathbb{R}^2$ and remove $\begin{pmatrix}0\\0\end{pmatrix}$

$\mathbb{R}\left\{\begin{pmatrix}-1\\-6\end{pmatrix}, \boxed{\begin{pmatrix}-2\\3\end{pmatrix}}, \cancel{\begin{pmatrix}0\\0\end{pmatrix}}, \begin{pmatrix}3\\0\end{pmatrix}\right\}$

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

Vector Space

- A real vector space is a set/collection of "*vectors*" together with the rules for closed under vector addition and closed under scalar multiplication by real numbers.*

*Strang, Gilbert. *Linear Algebra and Its Applications*. Cengage Learning, 2017.

Dimension and Basis of a Vector Space

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

$\boxed{\alpha_i \begin{pmatrix}1\\0\end{pmatrix} + \beta_i \begin{pmatrix}0\\1\end{pmatrix} \quad \alpha_i, \beta_i \in \mathbb{R}}$

Span of $\left\{ \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\}$

$\begin{pmatrix}5\\4\end{pmatrix} = \alpha_i \begin{pmatrix}1\\0\end{pmatrix} + \beta_i \begin{pmatrix}2\\0\end{pmatrix}$

$\left\{ \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix} \right\}$ $\dim(\mathbb{R}^2) = 2$

$\dim(\mathbb{R}^3) = 3$

Dimension of a Vector space - Every vector space has a dimension. Dimension is the number of basis vectors required to span the vector space.

Properties of Basis Vectors -

- Basis vectors has to be linearly independent.
- Basis vectors should span the vector space.

$\begin{bmatrix}1 & 2\\2 & 4\end{bmatrix}$ are col_1, col_2 linearly independent or linearly dependent

$col_2 = \alpha\, col_1$

$\boxed{col_j = \sum_{i=1, i\neq j}^{n} \alpha_i\, col_i}$

linearly dependent

$\begin{array}{ccc} col_1 & col_2 & col_3 \end{array}$

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

are these col's linearly dependent or independent

$col_3 = col_1 + col_2$

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

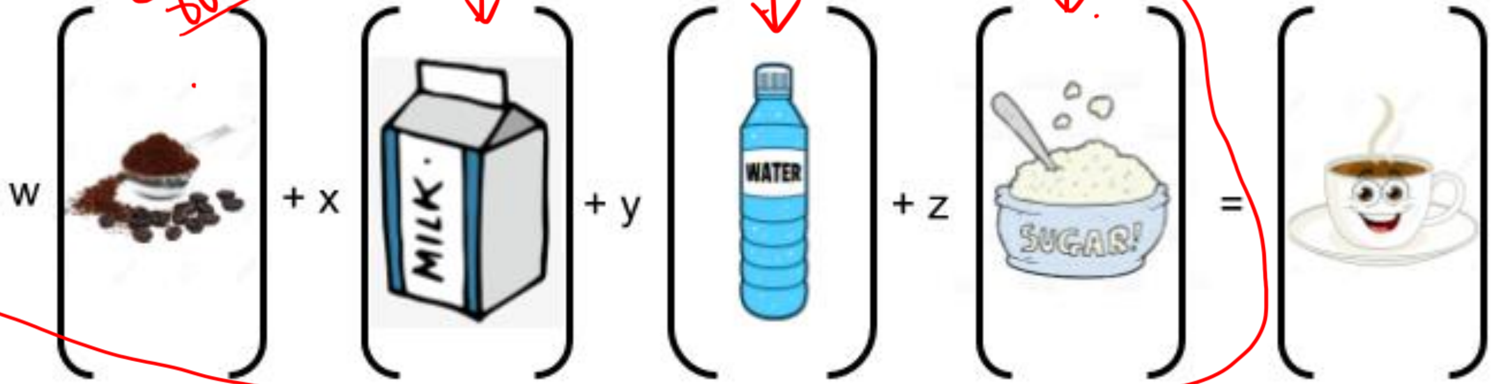
Dimension and Basis of a Coffee Space

- **Linear Independence**
- **Span the space**

{ $\vec{\text{coffee powder}}$, $\vec{\text{milk}}$, $\vec{\text{water}}$, $\vec{\text{sugar}}$ }

Coffee Space- Vector Space

coffee powder

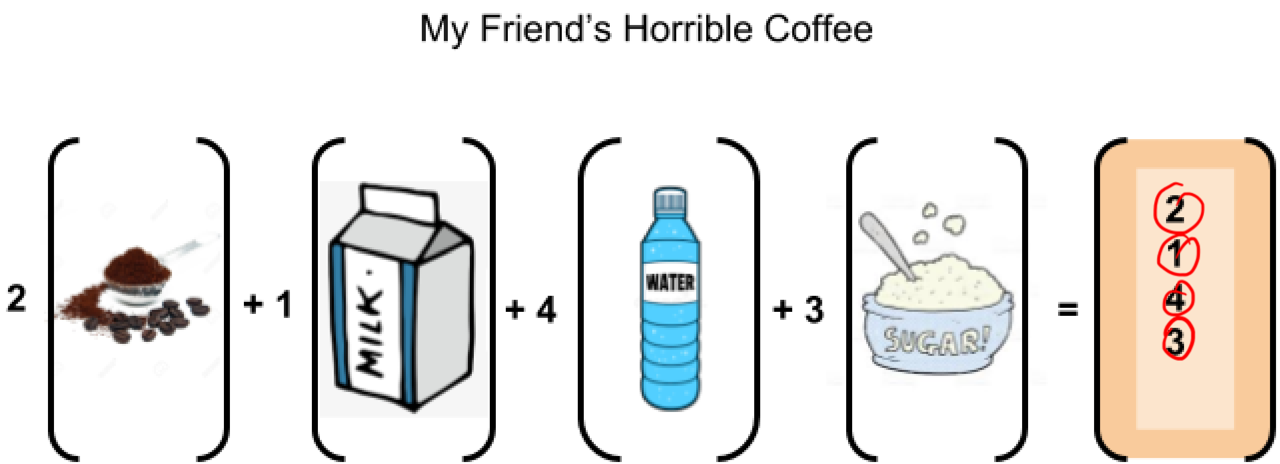


Coffee powder, milk, water and sugar are the basis vectors. Since there are only 4 basis vectors then coffee space has a dimension of 4.

Harikrishnan N B

My Friend’s Horrible Coffee

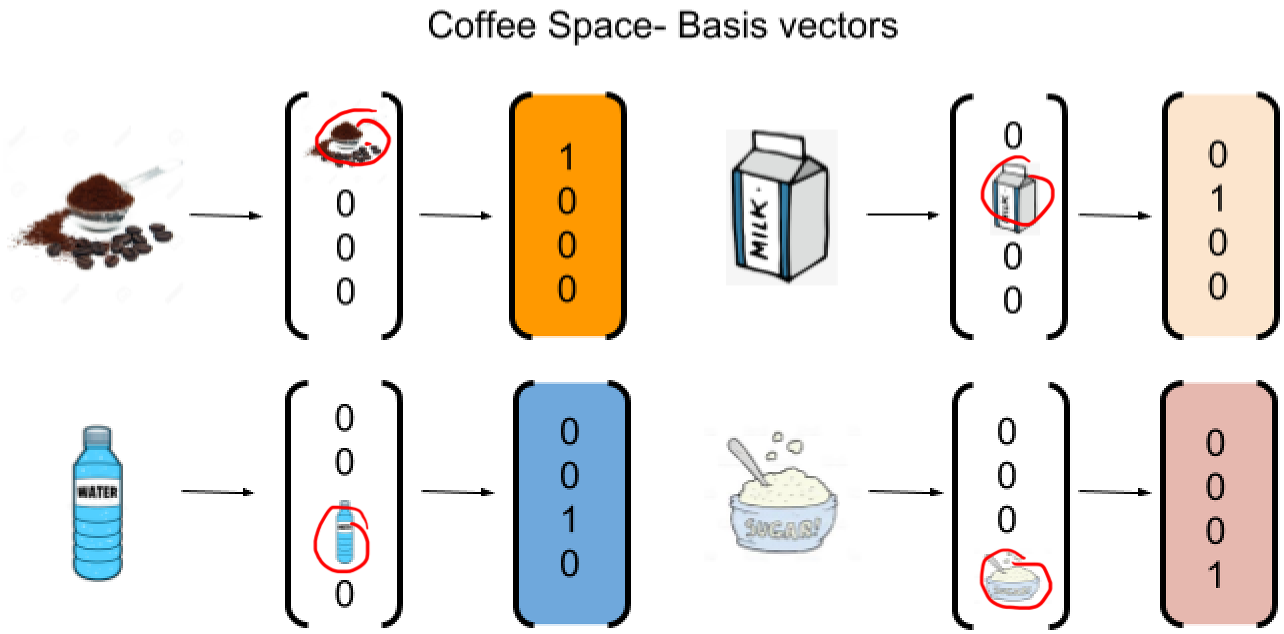
My Friend’s Horrible Coffee



My Friend's Horrible Coffee

$$2\begin{bmatrix}1\\0\\0\\0\end{bmatrix} + 1\begin{bmatrix}0\\1\\0\\0\end{bmatrix} + 4\begin{bmatrix}0\\0\\1\\0\end{bmatrix} + 3\begin{bmatrix}0\\0\\0\\1\end{bmatrix} = \begin{bmatrix}2\\1\\4\\3\end{bmatrix}$$

Visualizing Coffee Space Basis Vectors

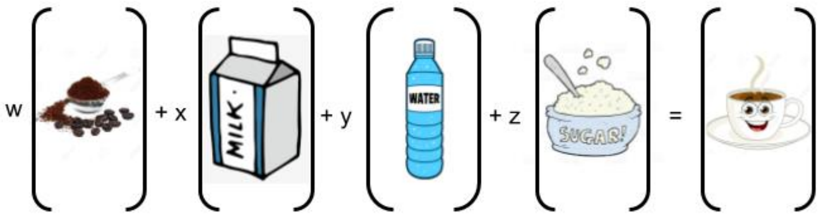


$$\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{bmatrix}5\\3\end{bmatrix} \Rightarrow x\begin{pmatrix}1\\0\end{pmatrix}+y\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}5\\3\end{pmatrix}$$

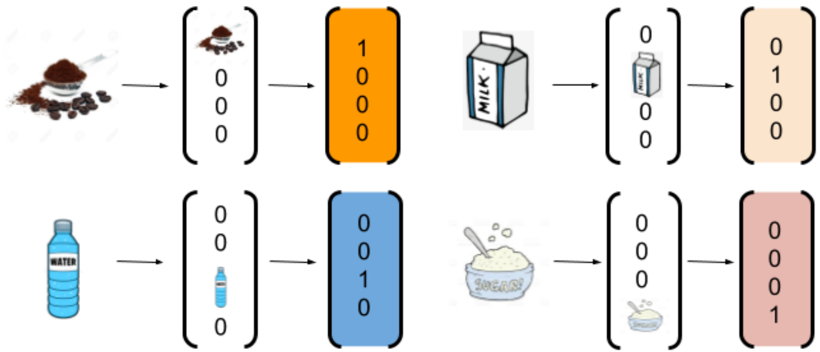
$$x+0y=5$$
$$0x+1y=3$$
$$x\begin{pmatrix}1\\0\end{pmatrix}+y\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}5\\3\end{pmatrix}$$

Matrix Multiplication - Visualization

Coffee Space- Vector Space



Coffee Space- Basis vectors



Ax = b

$$w\begin{bmatrix}1\\0\\0\\0\end{bmatrix} + x\begin{bmatrix}0\\1\\0\\0\end{bmatrix} + y\begin{bmatrix}0\\0\\1\\0\end{bmatrix} + z\begin{bmatrix}0\\0\\0\\1\end{bmatrix} = \begin{bmatrix}5\\6\\7\\8\end{bmatrix}$$

Ax = b

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix} = \begin{bmatrix}5\\6\\7\\8\end{bmatrix}$$

Column Space - Visualization

$C(A) \Rightarrow w\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + x\begin{pmatrix}0\\1\\0\\0\end{pmatrix} + y\begin{pmatrix}0\\0\\1\\0\end{pmatrix} + z\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$ $\quad w, x, y, z \in \mathbb{R}$

$\mathbb{R}^4$

$\dim(C(A)) = 4$

Ax = b

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix}$$

$\mathbb{R}^2$

$x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix} + z\begin{pmatrix}1\\1\end{pmatrix}$

$x, y, z \in \mathbb{R}^2$

$\begin{bmatrix} \begin{pmatrix}1\\0\end{pmatrix} & \begin{pmatrix}0\\1\end{pmatrix} & \begin{pmatrix}1\\1\end{pmatrix} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{pmatrix}9\\6\end{pmatrix}$

Ax = b $\quad \dim(C(A))$

$$w\begin{bmatrix}1\\0\\0\\0\end{bmatrix} + x\begin{bmatrix}0\\1\\0\\0\end{bmatrix} + y\begin{bmatrix}0\\0\\1\\0\end{bmatrix} + z\begin{bmatrix}0\\0\\0\\1\end{bmatrix} = \begin{bmatrix}5\\6\\7\\8\end{bmatrix}$$

Column Space of Matrix A - Column space of matrix A denoted as C(A) is the space spanned by the column vectors of A.

C(A)

$$span\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}$$

Dimension of C(A) = 4. Since 4 linearly independent vectors are there in the columns of matrix A. These vectors act as the basis and span the entire R^4.

Thinking

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \end{bmatrix}}_{A} \begin{pmatrix} x \\ y \\ z \\ w \\ k \end{pmatrix}$$

$C(A) = 2\begin{pmatrix}1\\2\end{pmatrix}, 3\begin{pmatrix}1\\2\end{pmatrix}, 4\begin{pmatrix}1\\2\end{pmatrix}, 5\begin{pmatrix}1\\2\end{pmatrix}$

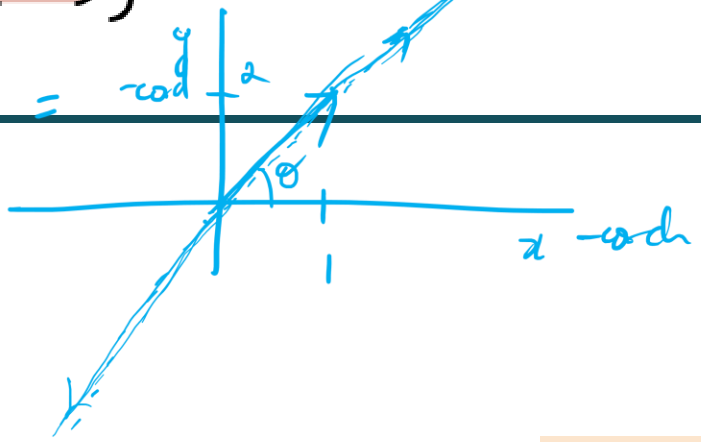
$C(A) = \text{span}\left\{ \begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}2\\4\end{pmatrix}, \begin{pmatrix}3\\6\end{pmatrix}, \begin{pmatrix}4\\8\end{pmatrix}, \begin{pmatrix}5\\10\end{pmatrix} \right\}$

$x\begin{pmatrix}1\\2\end{pmatrix} + y\begin{pmatrix}2\\4\end{pmatrix} + z\begin{pmatrix}3\\6\end{pmatrix} + w\begin{pmatrix}4\\8\end{pmatrix} + k\begin{pmatrix}5\\10\end{pmatrix}$

C(A)

Why $span\left\{ \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix} \right\}$ **can represent any point in R^4 ?**

$\dim(C(A)) =$



Harikrishnan N B

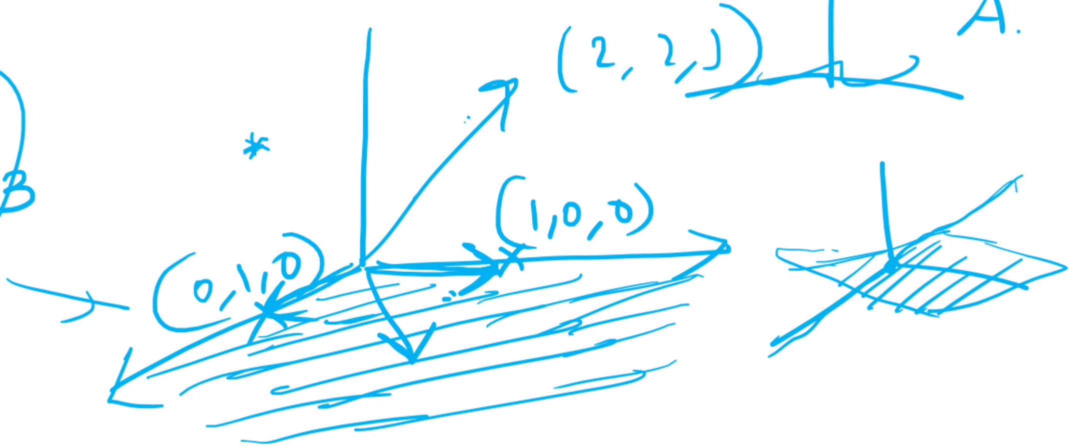
$Ax = b$ has a solution only b lies in the column space of A.

Can you see the Column Space?

$$\begin{array}{cc} col_1 & col_2 \\ \end{array}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w \\ x \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ \boxed{0} \end{bmatrix} \rightarrow B$$

Span{col_1, col_2}



$$w\begin{pmatrix} 1 \\ 0 \\ \boxed{0} \end{pmatrix} + x\begin{pmatrix} 0 \\ 1 \\ \boxed{0} \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} ?$$

$$\begin{pmatrix} 2 \\ 2 \\ 5 \end{pmatrix}$$

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

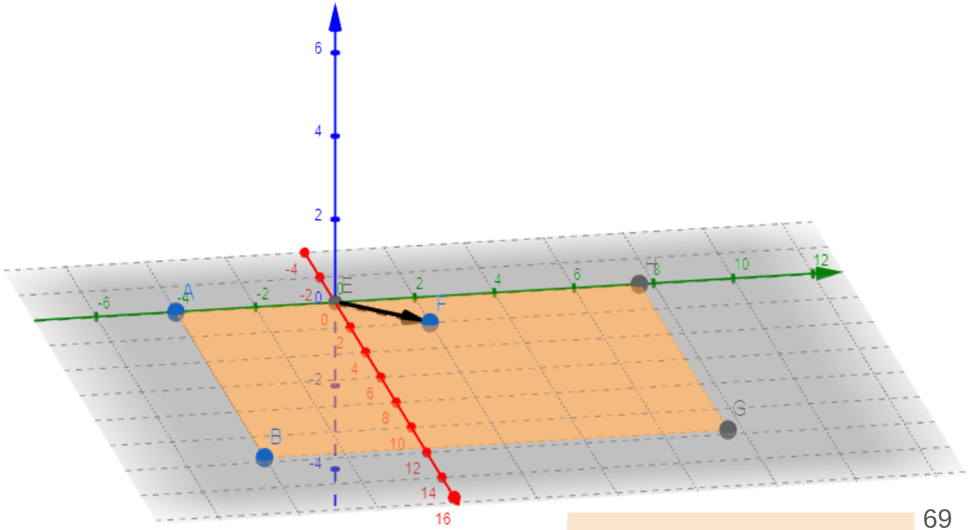
$$2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$$

Can you see the Column Space?

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} w \\ x \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \quad \longrightarrow \quad w \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$$

C(A)

$$span \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \longrightarrow$$



Linearly independent

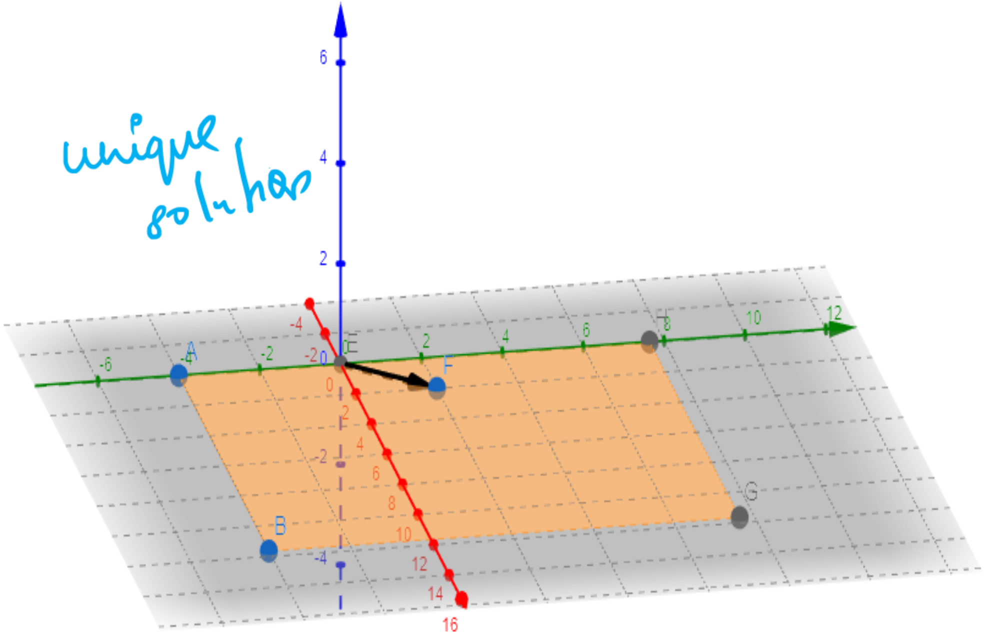
Some Observations !!!

$$\begin{bmatrix}1 & 0\\ 0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}w\\ x\end{bmatrix} = \begin{bmatrix}2\\ 2\\ 0\end{bmatrix}$$

C(A)

$$span\left\{\begin{bmatrix}1\\ 0\\ 0\end{bmatrix}, \begin{bmatrix}0\\ 1\\ 0\end{bmatrix}\right\}$$

unique solution



What is the dimension of Column space of Matrix A?

Will the basis vectors of C(A) span the entire 3-D space?

cols of A is linearly dependent & b lies in the col space of A then

What can you say about this?

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

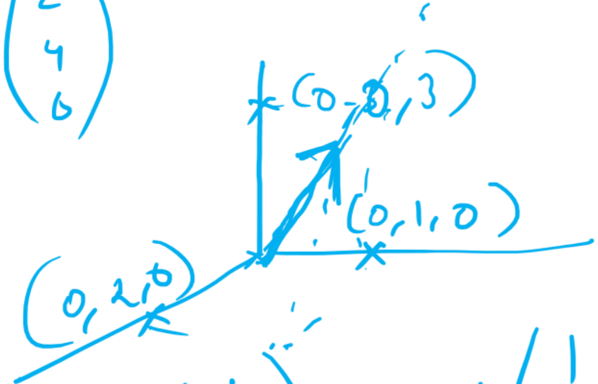
$$2\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$$

Identify the cols

$C(A) = \text{span}\{col_1, col_2\}$

$\Rightarrow w\, col_1 + x\, col_2$ where $w, x \in \mathbb{R}$

$\dim(C(A)) = 1$



$$w\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 2x\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

(1, 0)

(0, 1/2)

(2, -1)

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

What can you say about this?

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

$$w\begin{bmatrix}1\\ 2\\ 3\end{bmatrix}+2x\begin{bmatrix}1\\ 2\\ 3\end{bmatrix}=\begin{bmatrix}1\\ 2\\ 3\end{bmatrix}$$

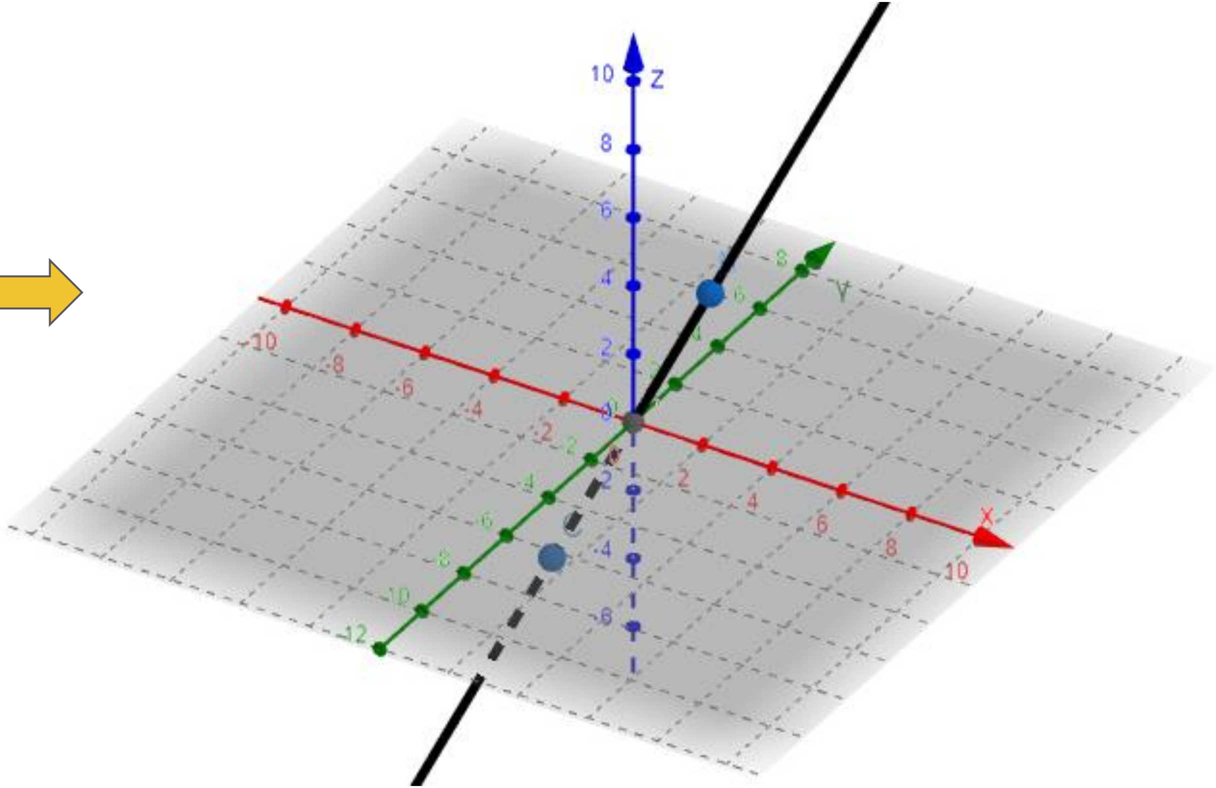
Column Space of Matrix A - Column space of matrix A denoted as C(A) is the space spanned by the column vectors of A.

$$C(A) = \mathit{span}\left\{\begin{bmatrix}1\\ 2\\ 3\end{bmatrix}\right\}$$

Dimension of C(A) = 1. Here Column space is a line passing through origin.

Do you see a Subspace ?

$$C(A) = span\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}\right\}$$



Harikrishnan N B

Is there anything Mysterious ?

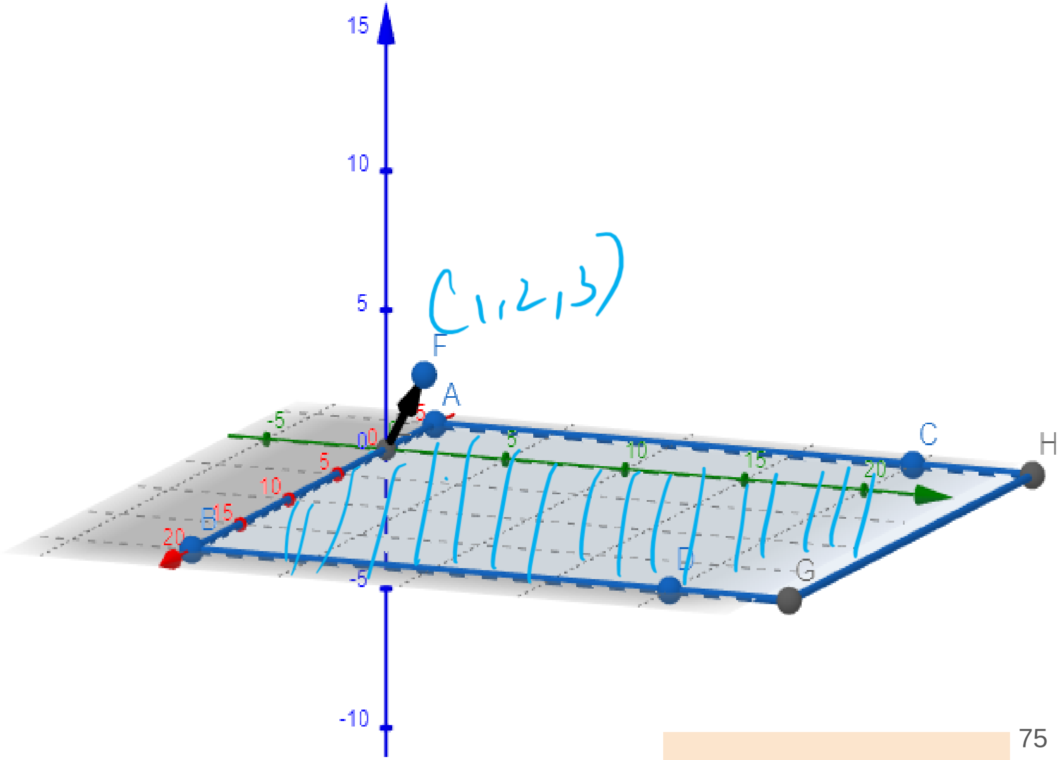
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Is there anything Mysterious ?

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

$$\Downarrow$$

$$w \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$



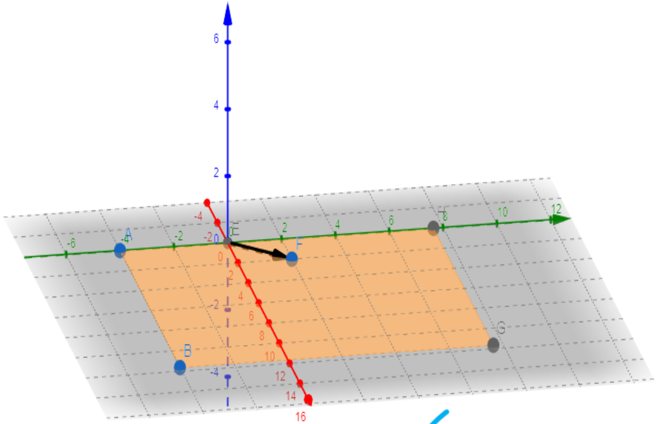
linearly depende-

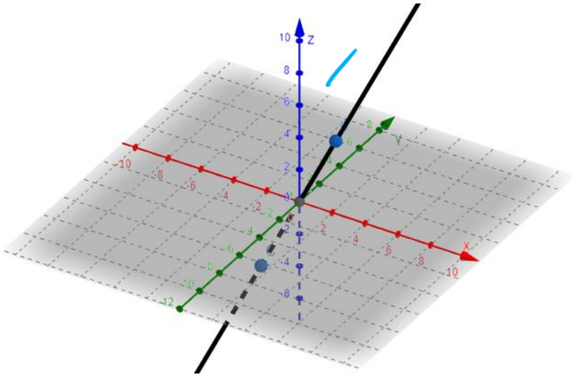
Solution to Ax = b

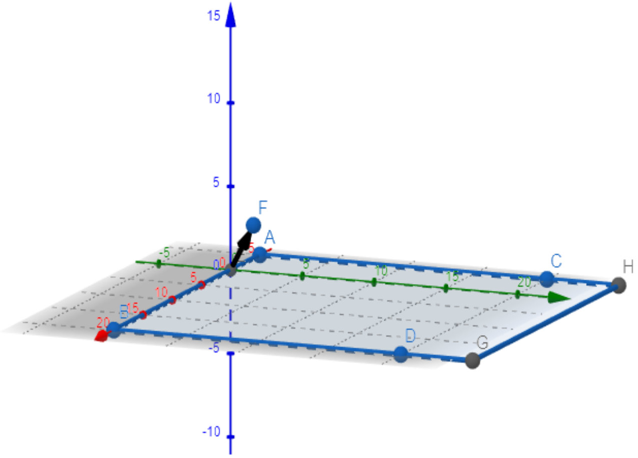
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w \\ x \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$$

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

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







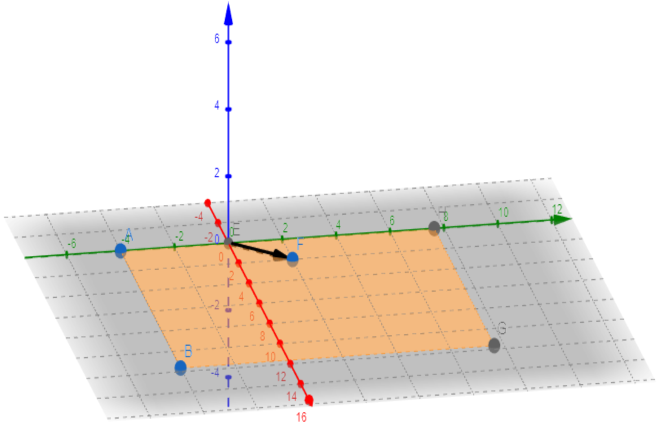
UNIQUE SOLUTION

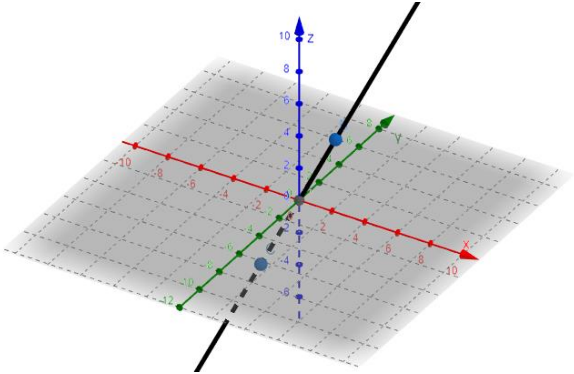
INFINITELY MANY SOLUTIONS

NO SOLUTION

So when does Ax = b have a Solution

Ax = b has solution when b lies in the column space of A or in other words b is a linear combination of column vectors of A.





- For unique solution and infinitely many solutions b lies in the column space of A.
- In the case of NO solution b does not lie in the column space of A.

UNIQUE SOLUTION

INFINITELY MANY SOLUTIONS

NO SOLUTION CASE 🙁

Can I find the best approximate solution ?

x_1 Jan 1 50

x_2 Feb 2 -30

3 100

4

$\text{error}^2 = p^2 + b^2$

$h^2 = p^2 + b^2$

unknown variables

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

$$A\vec{x} \neq \vec{b}$$

$$\begin{bmatrix} 50 \\ 30 \\ 100 \\ 20 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ \vdots & \vdots \\ 12 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} m \\ c \end{bmatrix}}_{\hat{x}} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

$A\vec{x} + \vec{e} = \vec{b}$

$\vec{e}$ is orthogonal to the col space

$\vec{e}$ is orthogonal span

$\vec{e} \perp col$

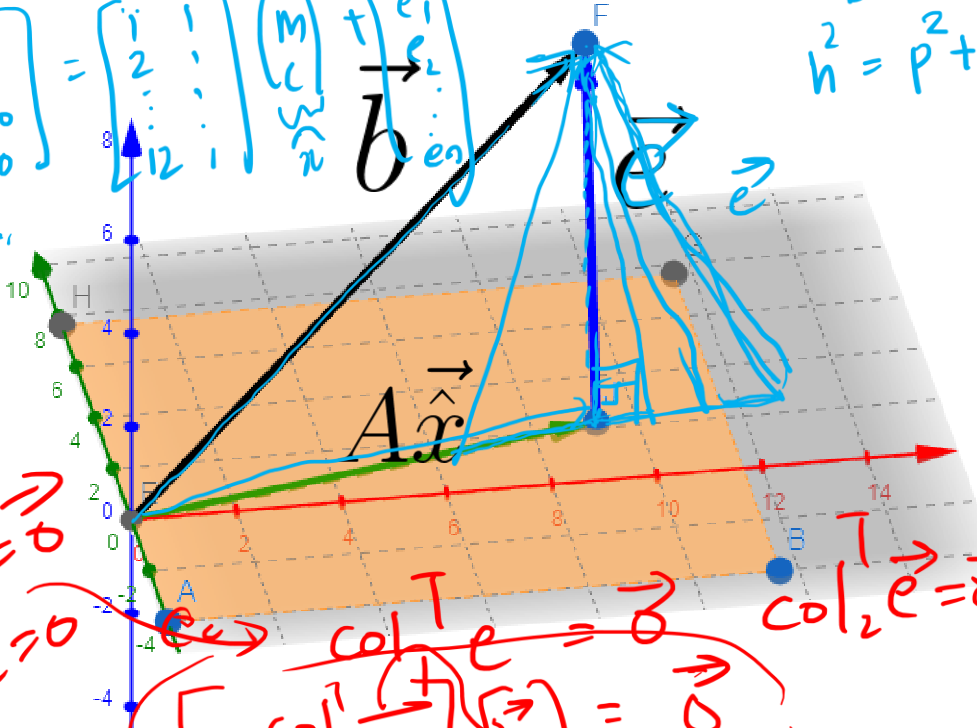
$col_1 \cdot \vec{e} = 0$

$col_2 \cdot \vec{e} = 0$

$col^T \vec{e} = \vec{0}$

$col_2^T \vec{e} = \vec{0}$

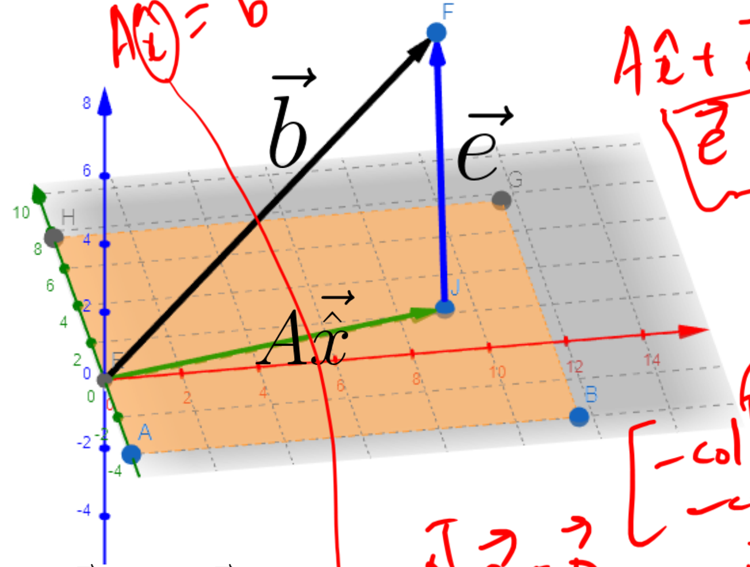
$$\begin{bmatrix} - col_1^T - \\ - col_2^T - \end{bmatrix} \begin{bmatrix} \vec{e} \end{bmatrix} = \vec{0}$$



No solution case - Visualization

$A\hat{x} \neq \vec{b}$

$A\hat{x} = b$



$A\hat{x} + \vec{e} = \vec{b}$

$\boxed{\vec{e} = \vec{b} - A\hat{x}}$

$$A\hat{\vec{x}} + \vec{e} = \vec{b}$$

$B^{-1}B = I$

$(A^TA)^{-1}(A^TA) = I$

$$\begin{bmatrix} b_1 & b_2 \end{bmatrix} \hat{\vec{x}} + \vec{e} = \vec{b}$$

$\begin{bmatrix} -\mathrm{col}_1^T- \\ -\mathrm{col}_2^T- \end{bmatrix} \vec{e} = \vec{0}$

$$\boxed{A^T\vec{e} = \vec{0}}$$

$A^TA\hat{x} = A^T\vec{b}$

$A^T\vec{e} = \vec{0}$

$A^T(\vec{b} - A\hat{x}) = \vec{0}$

$A^Tb - A^TA\hat{x} = \vec{0}$

$$\begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} \vec{e} = \vec{0}$$

$\hat{x} =$

$(A^TA)^{-1}A^TA\hat{x} = (A^TA)^{-1}A^T\vec{b}$

$-A^TA\hat{x} = -A^T\vec{b}$

$$A\hat{\vec{x}} + \vec{e} = \vec{b}$$
$$\vec{e} = \vec{b} - A\hat{\vec{x}}$$
$$A^T\vec{e} = \vec{0}$$
$$A^T(\vec{b} - A\hat{\vec{x}}) = \vec{0}$$
$$A^T\vec{b} - A^TA\hat{\vec{x}} = \vec{0}$$
$$A^TA\hat{\vec{x}} = A^T\vec{b}$$
$$\vec{x} = (A^TA)^{-1}A^T\vec{b}$$

ORTHOGONAL

NO SOLUTION CASE

Can I find the best approximate solution ?

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad A\vec{x} \neq \vec{b}$$

$$A\vec{\hat{x}} + \vec{e} = \vec{b}$$

$$\vec{e} = \vec{b} - A\vec{\hat{x}}$$

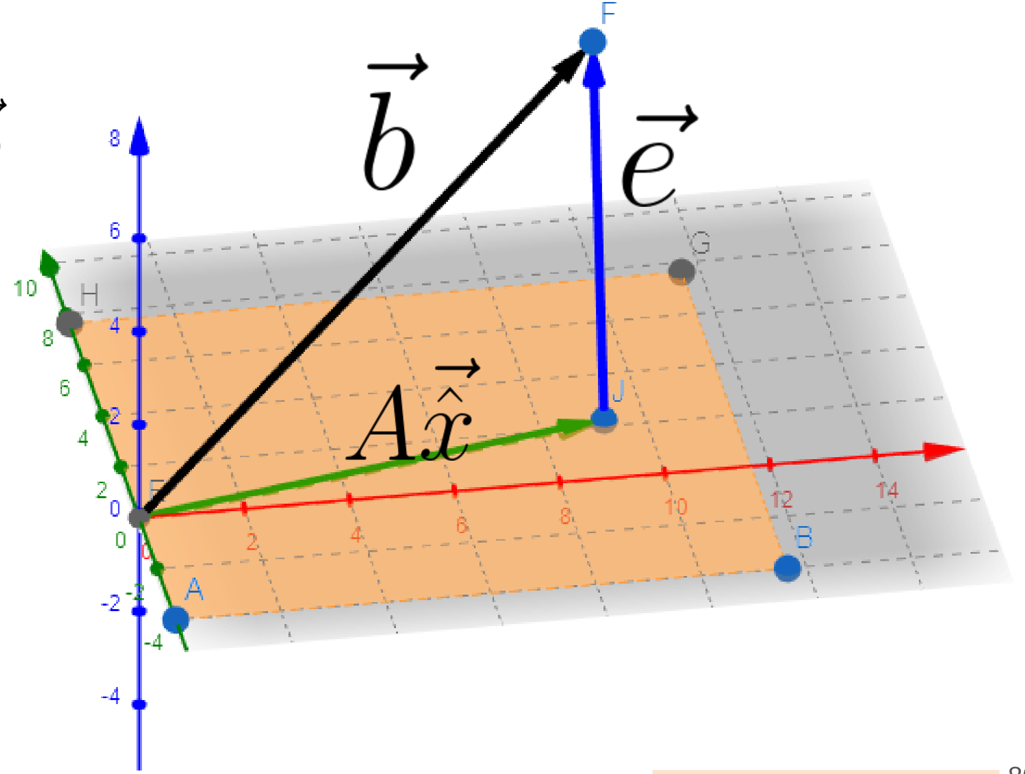
$$A^T\vec{e} = \vec{0}$$

$$A^T(\vec{b} - A\vec{\hat{x}}) = \vec{0}$$

$$A^T\vec{b} - A^T A\vec{\hat{x}} = \vec{0}$$

$$A^T A\vec{\hat{x}} = A^T\vec{b}$$

$$\vec{x} = (A^T A)^{-1}A^T\vec{b}$$



Linear Least Square Regression

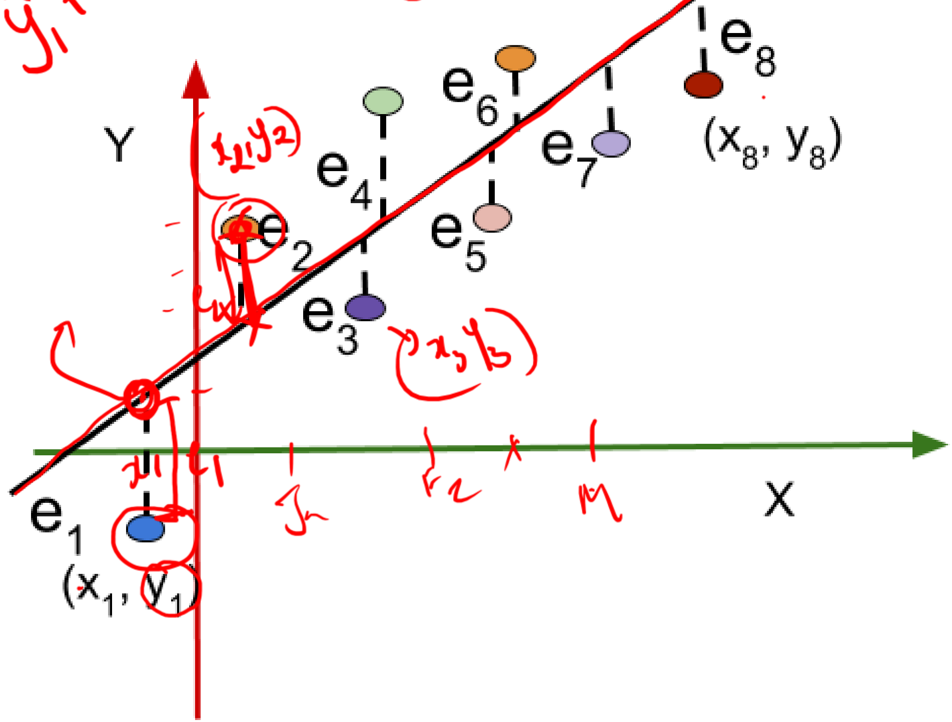
$\hat{y}_2 = \hat{y}_2 + e_2$

$\hat{y}_1 = mx_1 + c$

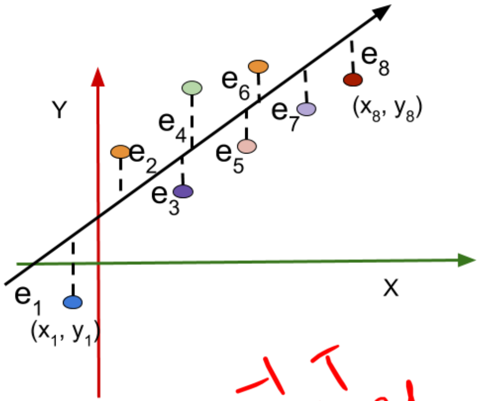
$\hat{y}_1 + e_1 = y_1$

$\hat{y}_2 = mx_2 + c$

$y = mx + c$



$$
\begin{aligned}
y_1 &= mx_1 + c + e_1 \\
y_2 &= mx_2 + c + e_2 \\
y_3 &= mx_3 + c + e_3 \\
y_4 &= mx_4 + c + e_4 \\
y_5 &= mx_5 + c + e_5 \\
y_6 &= mx_6 + c + e_6 \\
y_7 &= mx_7 + c + e_7 \\
y_8 &= mx_8 + c + e_8
\end{aligned}
$$



$$y_1 = mx_1 + c + e_1$$
$$y_2 = mx_2 + c + e_2$$
$$y_3 = mx_3 + c + e_3$$
$$y_4 = mx_4 + c + e_4$$
$$y_5 = mx_5 + c + e_5$$
$$y_6 = mx_6 + c + e_6$$
$$y_7 = mx_7 + c + e_7$$
$$y_8 = mx_8 + c + e_8$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} = \underbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \\ x_5 & 1 \\ x_6 & 1 \\ x_7 & 1 \\ x_8 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} m \\ c \end{bmatrix}}_{\hat{x}} + \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \end{bmatrix}}_{\vec{e}}$$

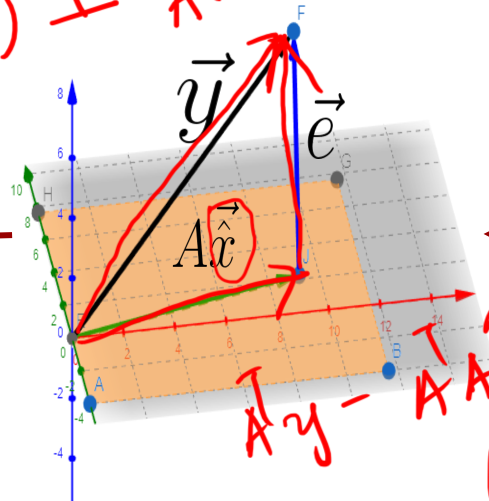
unknowns

$\hat{x} = (A^T A)^{-1} A^T y$

$\vec{e} = (y - A\hat{x}) \perp A^T$

$\vec{y} = A\hat{x} + \vec{e}$

$$\boxed{\vec{y} = A\hat{\vec{x}} + \vec{e}}$$



$$\boxed{\vec{x} = (A^T A)^{-1} A^T \vec{y}}$$

$A^T(y - A\hat{x}) = 0$

$A^T y - A^T A\hat{x} = 0$

$A^T A\hat{x} = A^T y$

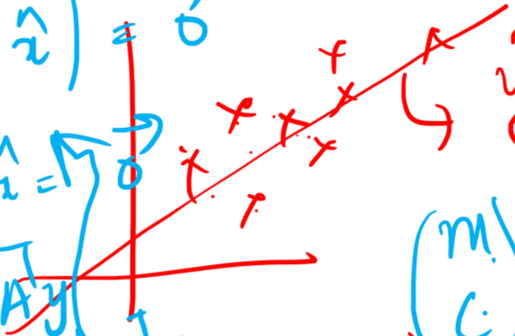
$\hat{x} = (A^T A)^{-1} A^T y$

$$A^T(y - A\hat{x}) = \vec{0}$$

$$A^Ty - A^TA\hat{x} = \vec{0}$$

$$A^TA\hat{x} = A^Ty$$

$$\hat{x} = (A^TA)^{-1}A^Ty$$



$$\hat{y} = \boxed{m}x + \boxed{c}$$

$$mx_i + c$$

$$\begin{pmatrix} m \\ c \end{pmatrix}$$

$$y_1 = \hat{y}_1 + e_1 = mx_1 + c + e_1$$

$$y_2 = \hat{y}_2 + e_2 = mx_2 + c + e_2$$

$$\vdots$$

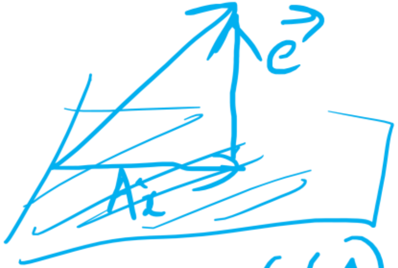
$$y_n = \hat{y}_n + e_n = mx_n + c + e_n$$

$$y \neq A\hat{x}$$

$$\boxed{\vec{y} = A\hat{x} + \vec{e}}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

$$\vec{y} = A\hat{x} + \vec{e}$$



$$\vec{e} \perp C(A)$$

$$A^T\vec{e} = \vec{0}$$

$$\begin{bmatrix} - \mathrm{col}_1^T - \\ - \mathrm{col}_2^T - \end{bmatrix} \begin{bmatrix} \vec{e} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathrm{col}_1 \cdot \vec{e} = 0$$

$$\mathrm{col}_1^T \vec{e} = 0 \qquad \mathrm{col}_2^T \cdot \vec{e} = 0$$

What happens in this case ? - Code it

$$\boxed{C(A) \perp \vec{e}}$$

$$A^T = \begin{bmatrix} -a_1^T- \\ -a_2^T- \\ \vdots \\ -a_n^T- \end{bmatrix} \quad A^T\vec{e} = \vec{0}$$

$$A = \begin{bmatrix} | & | & | & & | \\ a_1 & a_2 & a_3 & \cdots & a_n \\ | & | & | & & | \end{bmatrix}$$

$$a_i \cdot \vec{e} = \vec{0}$$

$$a_i^T e = 0$$

$$A^T \cdot \vec{e} =$$

$$\begin{bmatrix} -a_1^T- \\ -a_2^T- \\ \vdots \\ -a_n^T- \end{bmatrix} \begin{bmatrix} | \\ \vec{e} \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} = \begin{bmatrix} x_1 & x_1^2 & 1 \\ x_2 & x_2^2 & 1 \\ x_3 & x_3^2 & 1 \\ x_4 & x_4^2 & 1 \\ x_5 & x_5^2 & 1 \\ x_6 & x_6^2 & 1 \\ x_7 & x_7^2 & 1 \\ x_8 & x_8^2 & 1 \end{bmatrix} \begin{bmatrix} m \\ p \\ c \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \end{bmatrix}$$

$$\vec{y} = A\vec{x} + \vec{e}$$

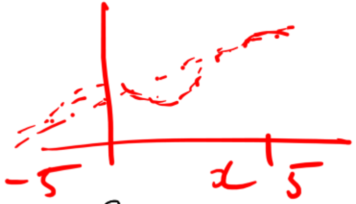
What change you will observe in the graph?

What happens when you add more higher order terms like x^3, x^4.. x^n?

$$\boxed{\vec{x} = (A^T A)^{-1} A^T \vec{y}}$$

$$\boxed{y\_data = x^3 + \sin x + \text{noise}}$$

step 1 $\quad$ (y_data vs x)



$$y\_data = \begin{bmatrix} y\_data(0) \\ y\_data(1) \\ \vdots \\ y\_data(n) \end{bmatrix} \qquad x = \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(n) \end{bmatrix}$$

$$m x_0 + k x_0^2 + c$$

$$y\_data = \underbrace{\begin{bmatrix} x_0 & x_0^2 & 1 \\ x_1 & x_1^2 & 1 \\ \vdots & \vdots & \vdots \\ x_n & x_n^2 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} m \\ k \\ c \end{bmatrix}}_{\hat{z}} + \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}}_{e}$$

$$\hat{z} = (A^T A)^{-1} A^T y\_data$$

$$\hat{z} \rightarrow \begin{pmatrix} m \\ c \end{pmatrix}$$

Applications of Least Squares in Signal Processing

- Linear/ Non-linear Prediction
- Denoising
- Deconvolution
- System Identification
- Estimating Missing Data

[Link to Ivan Selesnick's Tutorial]

LEAST SQUARES WITH EXAMPLES IN SIGNAL PROCESSING*

Ivan Selesnick

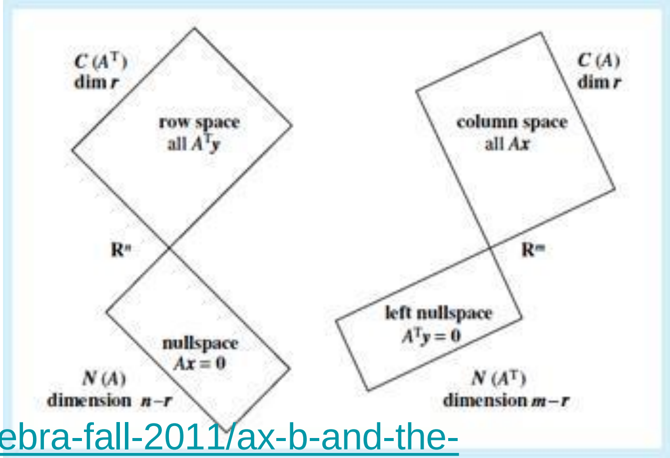
Four Fundamental Subspaces

- **Column Space**
- **Left Null Space**
- **Row Space**
- **Right Null Space**

Harikrishnan N B

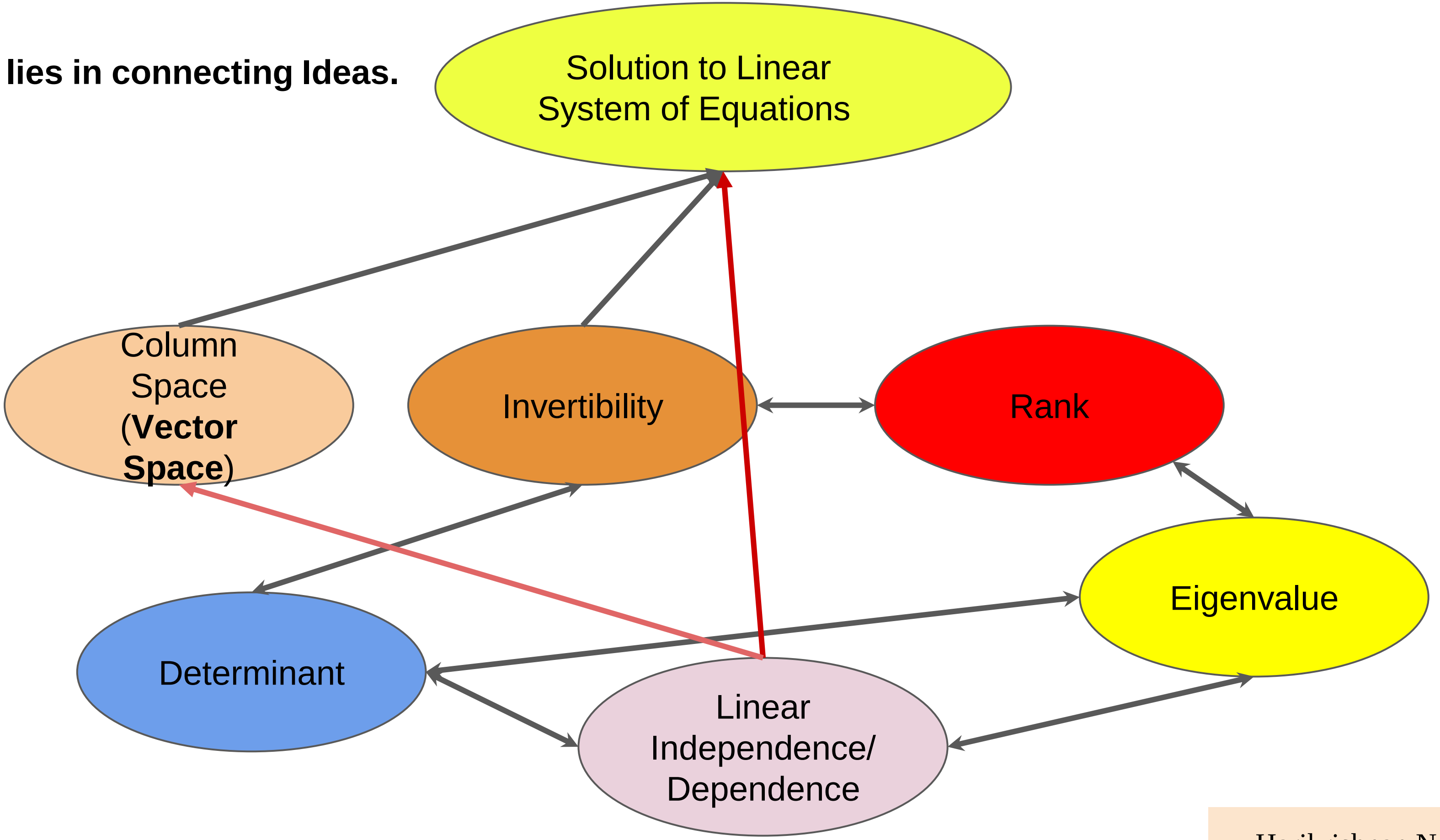
Fundamental Theorem of Linear Algebra

- Column space and Row space both have dimension r (rank).
- The Right Null Space have dimension *n-r* and the left null space has dimension *m-r*.
- Right Null Space is the orthogonal complement of the row space.
- Left Null Space is the orthogonal complement of the column space



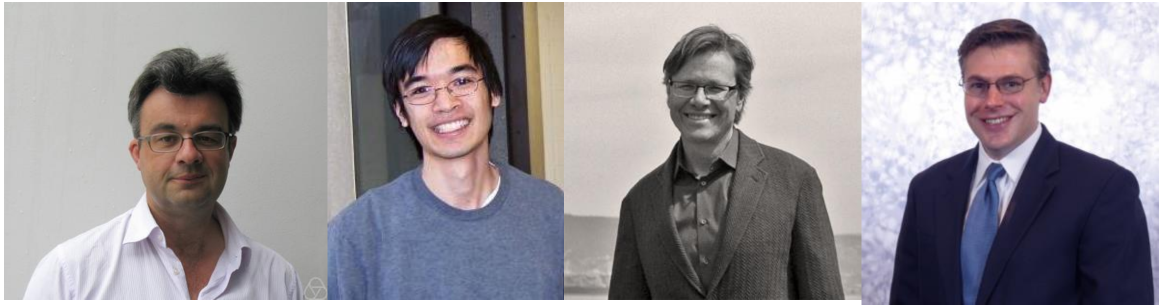
Source: https://ocw.aprende.org/courses/mathematics/18-06sc-linear-algebra-fall-2011/ax-b-and-the-four-subspaces/

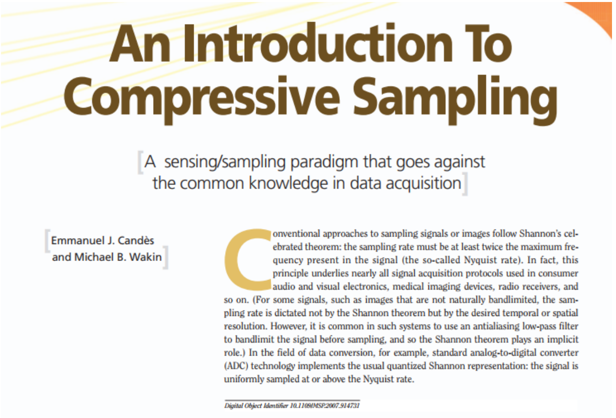
Beauty lies in connecting Ideas.



Do we know everything about *y=Ax* ?

Emmanuel Candès Terence Tao David Donoho Justin Romberg





An Introduction To Compressive Sampling

[A sensing/sampling paradigm that goes against the common knowledge in data acquisition]

[Emmanuel J. Candès and Michael B. Wakin]

Conventional approaches to sampling signals or images follow Shannon's celebrated theorem: the sampling rate must be at least twice the maximum frequency present in the signal (the so-called Nyquist rate). In fact, this principle underlies nearly all signal acquisition protocols used in consumer audio and visual electronics, medical imaging devices, radio receivers, and so on. (For some signals, such as images that are not naturally bandlimited, the sampling rate is dictated not by the Shannon theorem but by the desired temporal or spatial resolution. However, it is common in such systems to use an antialiasing low-pass filter to bandlimit the signal before sampling, and so the Shannon theorem plays an implicit role.) In the field of data conversion, for example, standard analog-to-digital converter (ADC) technology implements the usual quantized Shannon representation: the signal is uniformly sampled at or above the Nyquist rate.

Digital Object Identifier 10.1109/MSP.2007.914731

Near Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?

Emmanuel Candes† and Terence Tao♯

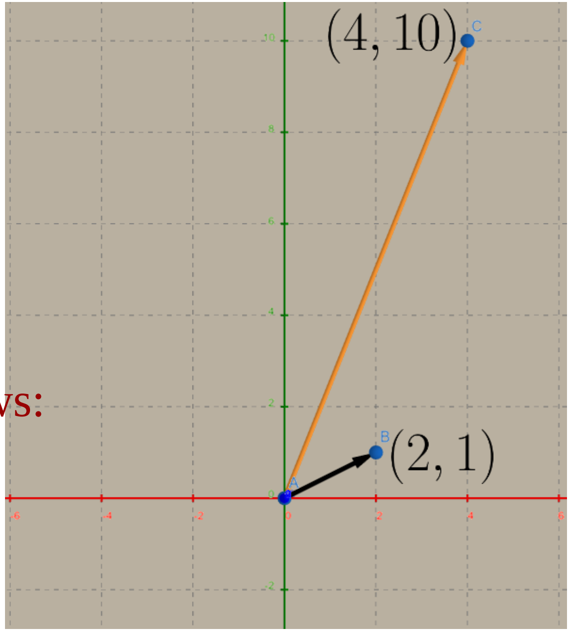
† Applied and Computational Mathematics, Caltech, Pasadena, CA 91125

♯ Department of Mathematics, University of California, Los Angeles, CA 90095

Matrix Vector Multiplication as a Transformation

Intuition for Matrix vector multiplication for Square Matrices

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \end{bmatrix}$$



Matrix(Square Matrix) vector multiplication can be seen as follows:

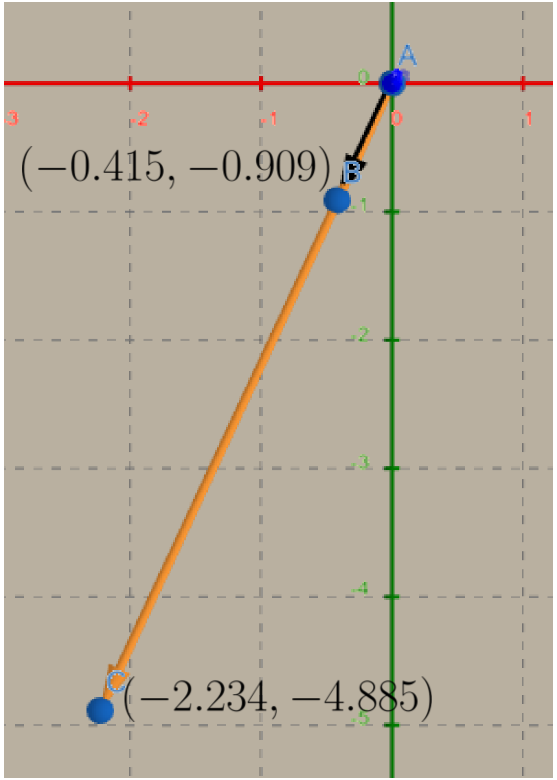
- Rotation
- Stretching or Shrinking

Special Vectors

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -0.415 \\ -0.909 \end{bmatrix} = \begin{bmatrix} -2.234 \\ -4.885 \end{bmatrix} = 5.372 \begin{bmatrix} -0.415 \\ -0.909 \end{bmatrix}$$

$$A\vec{x} \qquad\qquad \lambda\vec{x}$$

$$A\vec{x} = \lambda\vec{x}$$



1. Direction of $\vec{x}$ is unchanged. (No rotation)
2. Only the magnitude is scaled by a factor λ
3. $\vec{x}$ - **eigenvector of matrix A**
4. λ - **eigenvalue of matrix A**

$$A\vec{x} = \lambda\vec{x}$$

$$A\vec{x} = \lambda\vec{x}$$

Eigenvalues and Eigenvectors

- For an n x n square matrix A, there are ‘n’ eigenvalues and ‘n’ eigenvectors. Let x_1, x_2, x_3,....x_n be the ‘n’ eigenvectors and $\lambda_1, \lambda_2, \lambda_3, \ldots\lambda_n$, are the corresponding eigenvalues.

$$A\vec{x_1} = \lambda_1\vec{x_1}$$
$$A\vec{x_2} = \lambda_2\vec{x_2}$$
$$A\vec{x_3} = \lambda_3\vec{x_3}$$
$$.$$
$$.$$
$$A\vec{x_n} = \lambda_n\vec{x_n}$$

$$X = \begin{bmatrix} | & | & | & & | \\ \vec{x_1} & \vec{x_2} & \vec{x_3} & \cdots & \vec{x_n} \\ | & | & | & & | \end{bmatrix}$$

Very Very Important Part

- For an n x n square matrix A, there are ‘n’ eigenvalues and ‘n’ eigenvectors. Let x_1, x_2, x_3,....x_n be the ‘n’ eigenvectors and $\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_n$, are the corresponding eigenvalues.

$$AX = \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} | & | & | & & | \\ \vec{x}_1 & \vec{x}_2 & \vec{x}_3 & \cdots & \vec{x}_n \\ | & | & | & & | \end{bmatrix}$$

Spectral Decomposition

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}\begin{bmatrix} | & | & | & & | \\ \vec{x}_1 & \vec{x}_2 & \vec{x}_3 & \cdots & \vec{x}_n \\ | & | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & | & & | \\ A\vec{x}_1 & A\vec{x}_2 & A\vec{x}_3 & \cdots & A\vec{x}_n \\ | & | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & | & & | \\ \lambda_1\vec{x}_1 & \lambda_2\vec{x}_2 & \lambda_3\vec{x}_3 & \cdots & \lambda_n\vec{x}_n \\ | & | & | & & | \end{bmatrix}$$

$$\begin{bmatrix} | & | & | & & | \\ \lambda_1\vec{x}_1 & \lambda_2\vec{x}_2 & \lambda_3\vec{x}_3 & \cdots & \lambda_n\vec{x}_n \\ | & | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & | & & | \\ \vec{x}_1 & \vec{x}_2 & \vec{x}_3 & \cdots & \vec{x}_n \\ | & | & | & & | \end{bmatrix}\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Spectral Decomposition

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} | & | & | & & | \\ \vec{x}_1 & \vec{x}_2 & \vec{x}_3 & \cdots & \vec{x}_n \\ | & | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & | & & | \\ \vec{x}_1 & \vec{x}_2 & \vec{x}_3 & \cdots & \vec{x}_n \\ | & | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$AX = X\Lambda$$

$$AXX^{-1} = X\Lambda X^{-1}$$

$$AI = X\Lambda X^{-1}$$

$$A = X\Lambda X^{-1}$$

Harikrishnan N B

Practical Challenges and Important Points

When can we apply $A = X\Lambda X^{-1}$?

- A should be a square matrix
- When A has 'n' linearly independent eigenvectors, then X^{-1} always exist.

What happens when A is Symmetric ($A^T = A$)?

- The eigenvectors of a symmetric matrix A can be chosen as ORTHONORMAL. So in this case X is orthonormal.
- For an ORTHONORMAL matrix X, the inverse is its transpose $X^{-1} = X^T$

$$A = X\Lambda X^{-1}$$

$$A = X\Lambda X^T$$

Harikrishnan N B

Practical Challenges

What if A is not a square matrix?

- We cannot apply Spectral Decomposition.

Don't Worry!!!

Singular Value Decomposition works for any Matrix.

A Few more steps to PCA

What all minimum can we say about this data?

X	1	2	3	4	5
Y	1	5	4	6	7

X, Y are the features

What all minimum can we say about this data?

Mean(X) = $$\frac{1}{N}\sum_{i=1}^{N} x_i$$

Variance(X) = $$\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu_x)^2$$

Cov(X,Y) = $$\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu_x)(y_i-\mu_y)$$

X	Y	var(X)	var(Y)	cov(X,Y)	cov(Y,X)
1	1				
2	5				
3	4				
4	6				
5	7				
Mean (X)	Mean(Y)	var(X)	var(Y)	cov(X,Y)	cov(Y,X)
3.0	**4.6**	**2.5**	**5.3**	**2.25**	**2.25**

Variance- Covariance Matrix

$$\begin{array}{cc} & \begin{array}{cc} X & Y \end{array} \\ \begin{array}{c} X \\ Y \end{array} & \begin{bmatrix} var(X) & cov(X,Y) \\ cov(Y,X) & var(Y) \end{bmatrix} \end{array}$$

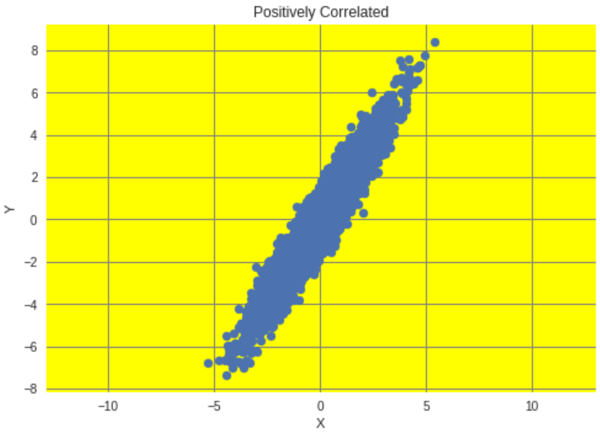
Recall the properties of a symmetric matrix!!!

- **Variance - Covariance Matrix is symmetric. cov(X,Y) = cov(Y,X)**
- **The diagonal entries represents variance**
- **The off- diagonal entries represents the correlation of X and Y**

What does Variance - Covariance Matrix signifies?

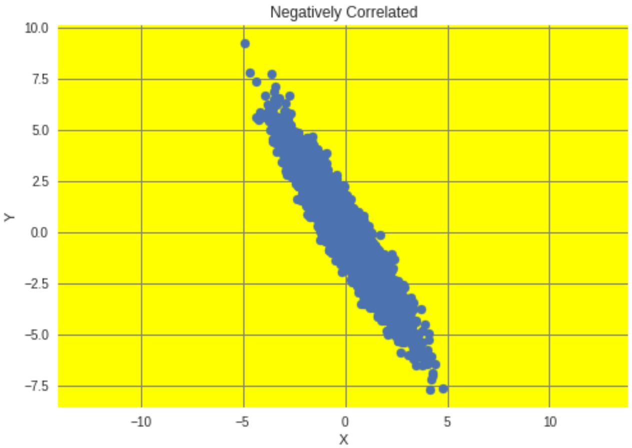
Case I

$$\begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$



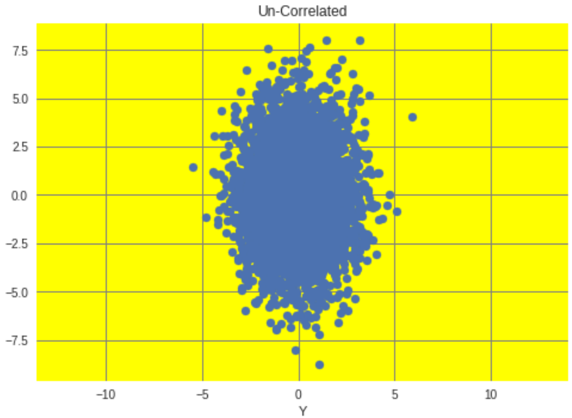
Case II

$$\begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}$$



Case III

$$\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$$

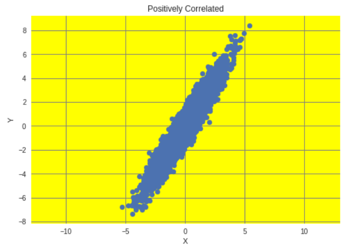


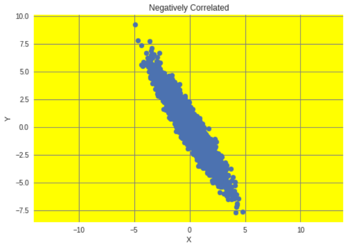
Note: In all cases mean is (0,0)

Harikrishnan N B

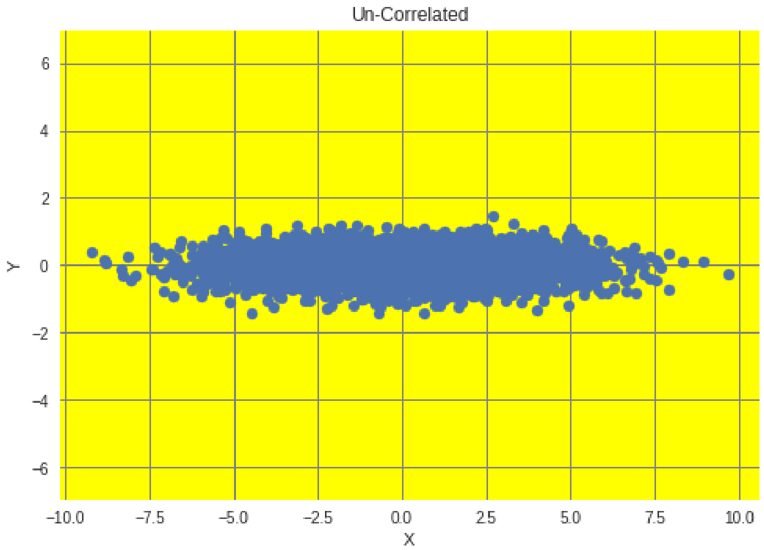
So what does PCA do ?

- Principal Component Analysis (PCA) makes the data **UNCORRELATED.**





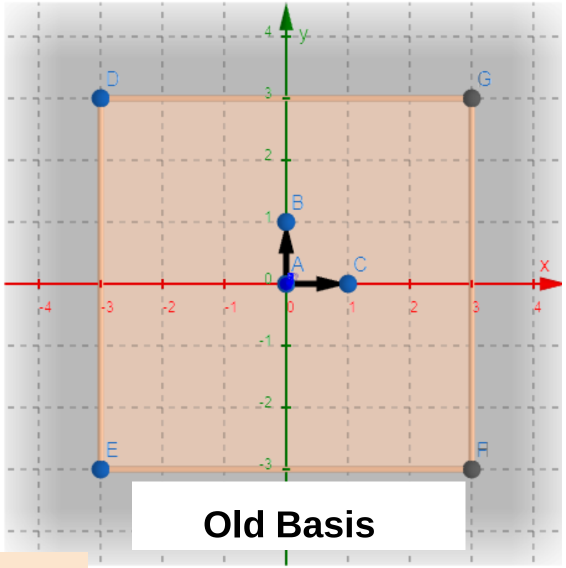
PCA achieves this by **Change of Basis**

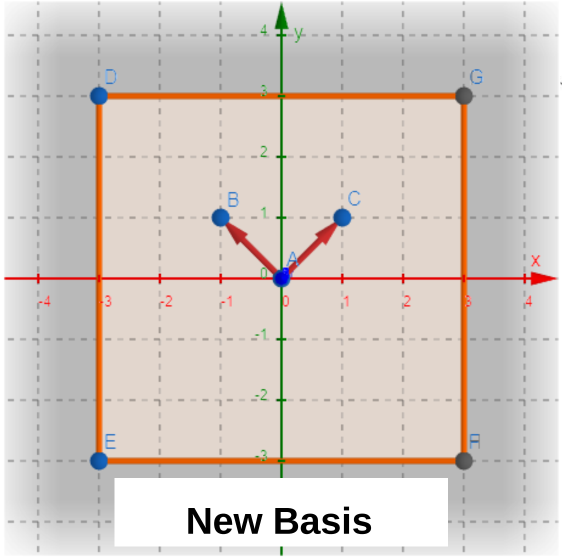


Change of Basis

My life is full of struggle.

IT depends on the way you look at life. Try to see life from a different **POINT OF VIEW**





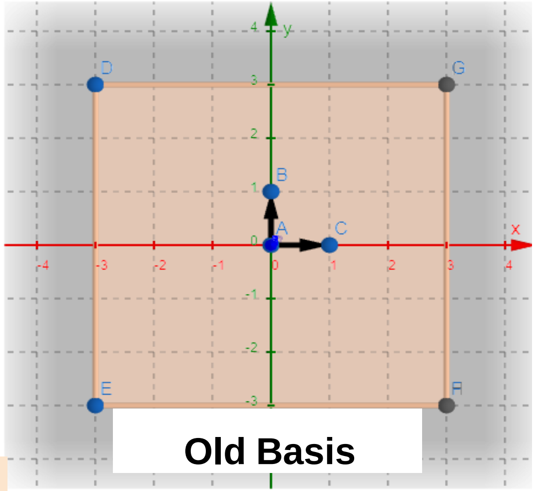
Harikrishnan N B

Recall

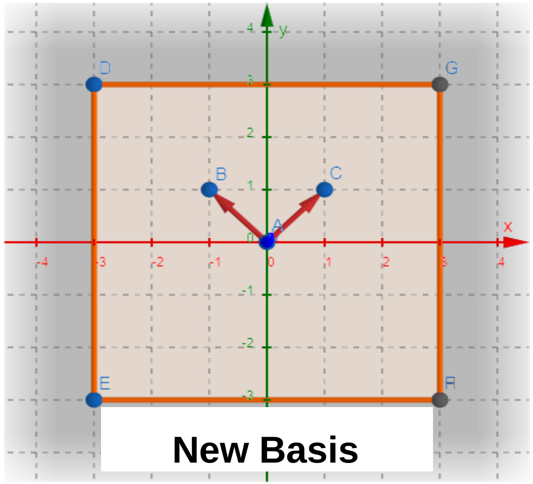
Dimension of a Vector space - Every vector space has a dimension. Dimension is the number of basis vectors required to span the vector space.

Properties of Basis Vectors -

- Basis vectors has to be linearly independent.
- Basis vectors should span the vector space.



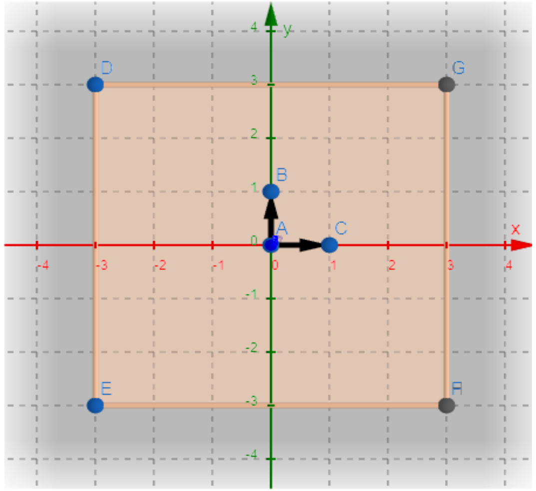
Old Basis



New Basis

Example of Change of Basis

To represent a point (2,3) in old basis and new basis- How to understand this?

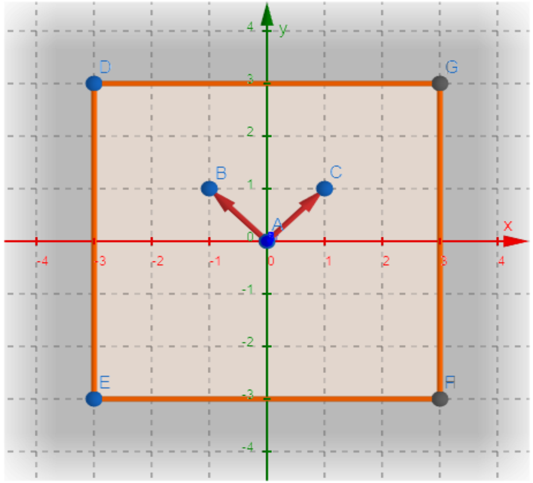


Old basis = $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$

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

$$\begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

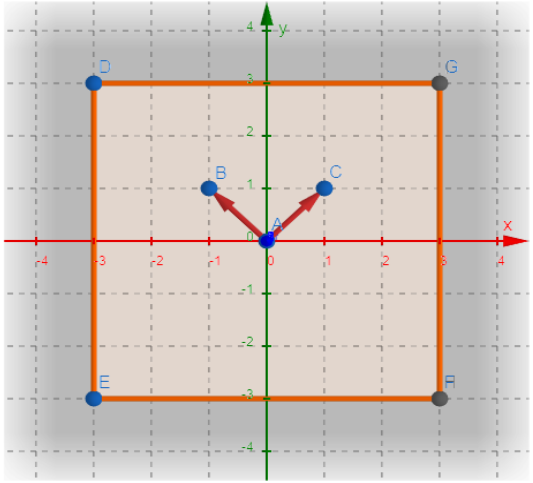
New Basis Representation



New Basis $= \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right\}$

$$x \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} + y \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

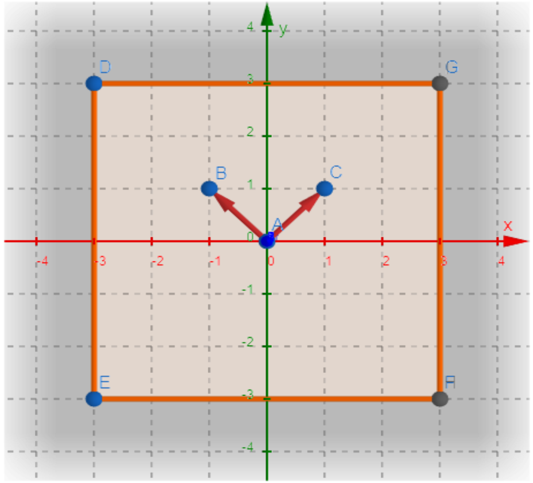
Finding x and y for representing (2,3) using new basis



$$x \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} + y \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

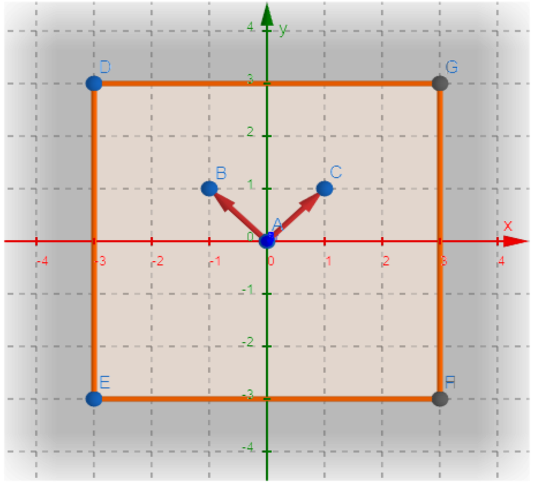


$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$$P\vec{x} = \vec{y}$$
$$P^{-1}P\vec{x} = P^{-1}\vec{y}$$
$$\vec{x} = P^{-1}\vec{y}$$

For **ORTHONORMAL MATRIX, $P^{-1} = P^T$**

In our case the matrix P is ORTHONORMAL



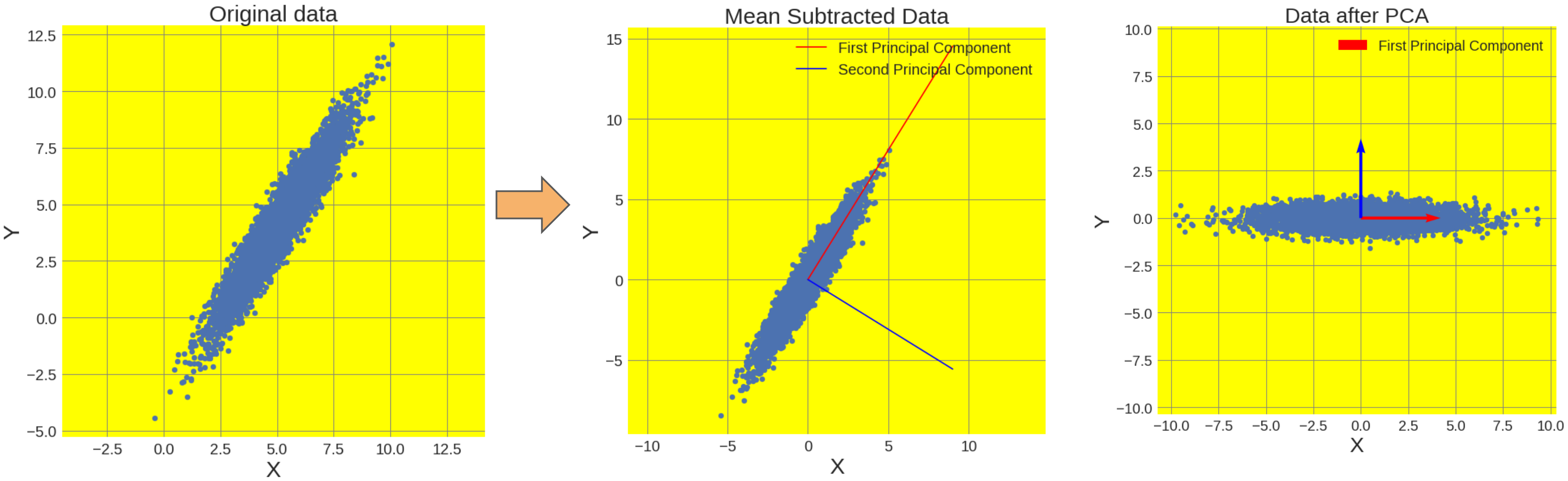
$$\vec{x} = P^{-1}\vec{y} = P^{T}\vec{y}$$

$$\begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{5}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\frac{5}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} + \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Harikrishnan N B

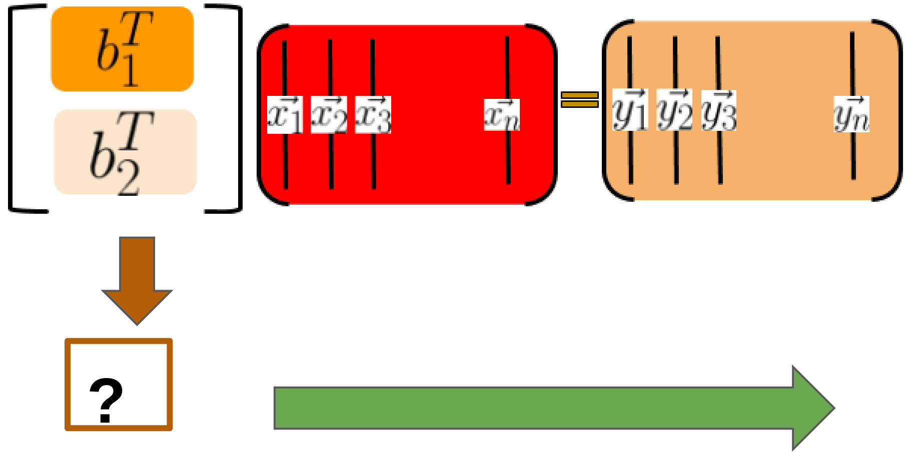
Steps in PCA



Change of Basis

$$\begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \vec{x}_3 & \cdots & \vec{x}_n \end{bmatrix} = \begin{bmatrix} \vec{y}_1 & \vec{y}_2 & \vec{y}_3 & \cdots & \vec{y}_n \end{bmatrix}$$

What should be the NEW BASIS so that DATA is UNCORRELATED?



Rows of matrix P are the **eigenvectors** of the **variance - covariance matrix** of the **mean subtracted data**

$$PX = Y$$
$$cov(Y) = cov(PX)$$
$$cov(PX) = \frac{1}{N-1}(PX)(PX)^T$$
$$cov(PX) = \frac{1}{N-1}PXX^TP^T$$
$$cov(PX) = P(\frac{1}{N-1}XX^T)P^T$$
$$cov(PX) = Pcov(X)P^T$$
$$cov(PX) = P(V\Lambda V^T)P^T$$
$$P = V^T$$
$$cov(PX) = \Lambda$$

Harikrishnan N B

Some words about PCA

- PCA is "an orthogonal linear transformation that transfers the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (*first principal component*), the second greatest variance lies on the second coordinate (*second principal component*), and so on."

Applications of PCA

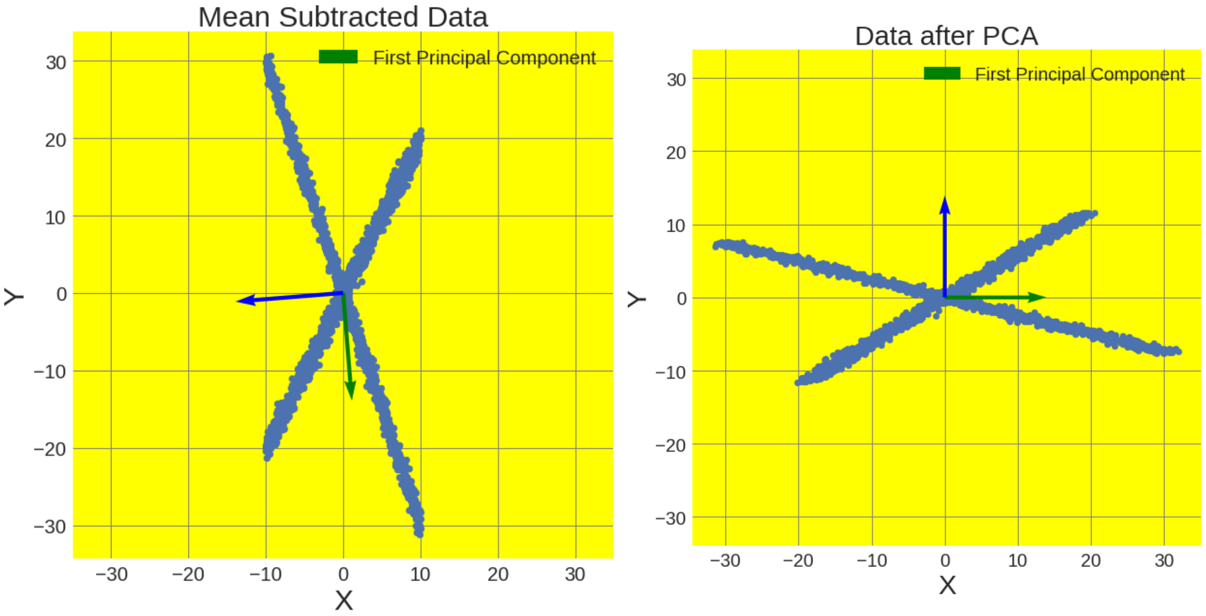
- Dimensionality Reduction
- Denoising
- Feature Extraction
- Image Compression
- EEG Analysis

Assumptions in PCA

- Linearity
- Large variance have important structure
- Principal components are orthogonal

When does PCA fail?

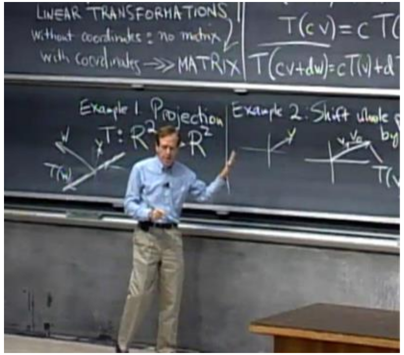
- Non-linearity
- Non-Gaussian
- Non-orthogonality

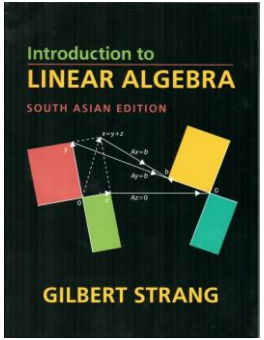


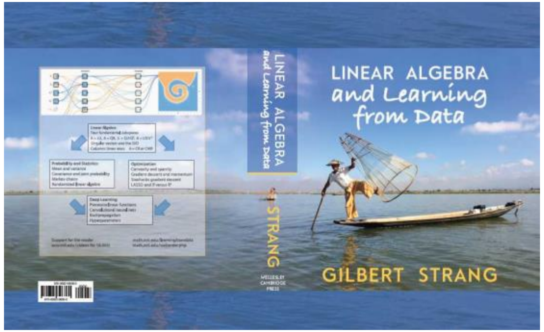
Ref: https://arxiv.org/abs/1404.1100

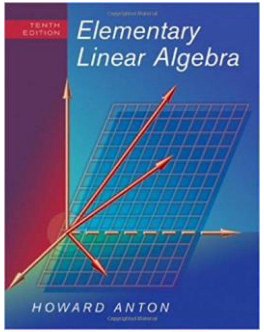
Interesting Materials

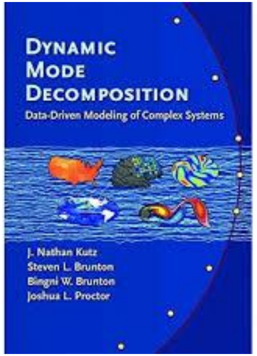
Prof. Gilbert Strang











Tutorial on PCA - (Click here)