23/12/23

Lecture-8

14/04/21

బొమ్మ పుట్ట రోజు

→ Continuing from last class of (Linear Algebraic transformation.)

→ We derived in last class we addressed the case of Linear Transformation that are not invertible

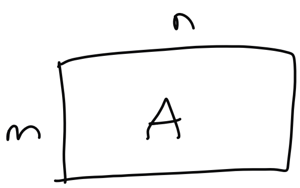
They are Non Injective & Non-Surjective cases

⇓

We made a derivation for Non-injective (but full Rank) case i.e for a Slanted Matrix 'A' i.e rows < columns

(Transformation Matrix)

$m < n$



& we derived the formula to calculate o/p PDF of o/p vector

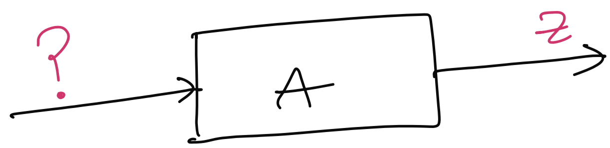
$$f_W(z) = \frac{1}{\sqrt{\det(AA^T)}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X\Big(\underbrace{A^T(AA^T)^{-1}}_{\text{pseudo inverse}} z + \underbrace{\sum_{j=0}^{n-m-1} \alpha_j e_j}_{\text{base vector of ker } A}\Big)\, d\alpha_0 \cdots d\alpha_{n-m-1}$$

given an input PDF

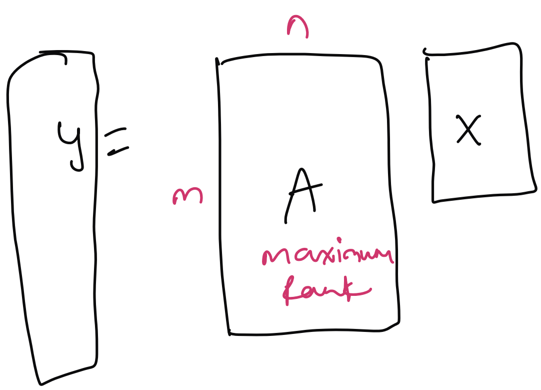
→ Now for Non-Surjective case

* Non surjective means sometimes we have a 'z' i.e instance of the o/p does not have any input (counter Image)

o/p cannot be produced by any input ~~and~~ This is Mathematically a problem but not physically. Since we are dealing with assigning probability to the o/p' we can just assign '0' and move on but let's give it a proper mathematical representation.

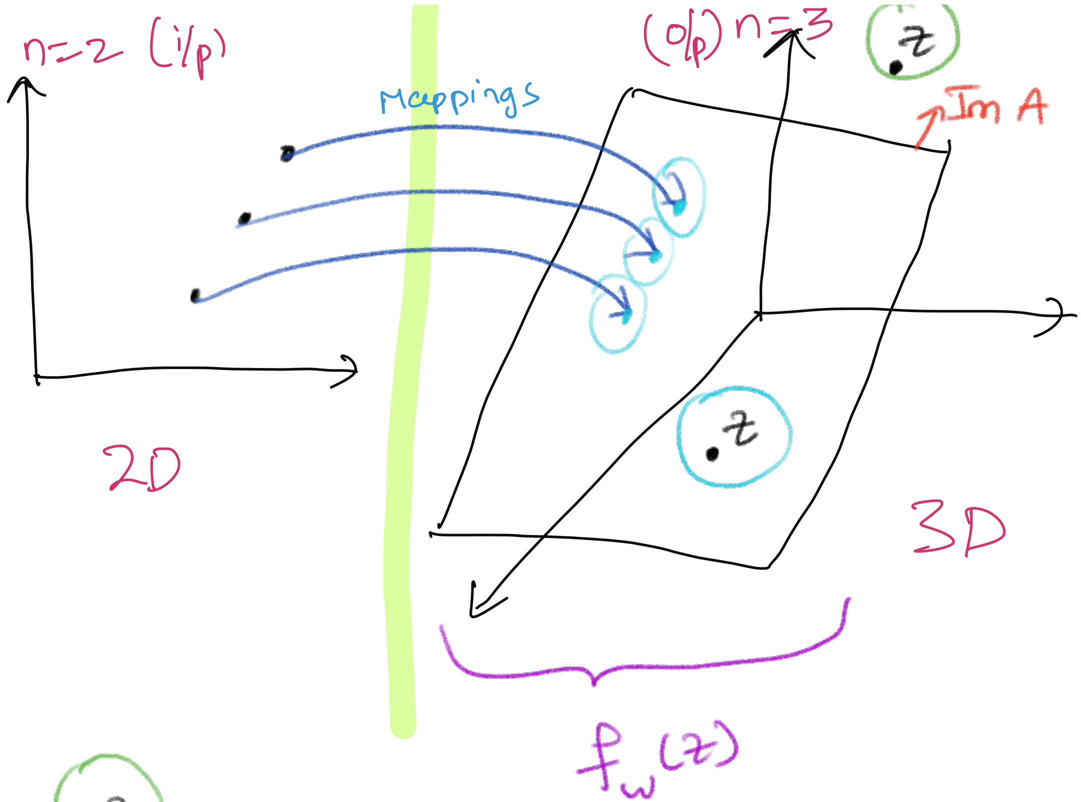


→ For certain 'z' we don't have counter images Then Matrix in this case will not be Slanted but a Tall Matrix.



$m > n$

• Implies The dimension of input x < The dimension of the output y



z out of the Im A

z Inside the Im A

Regrettable since the mapping is Linear i.e it maps straight things into straight things. so it cannot bend (or) fold anything.

∴ The mapping is from i/p plane to o/p plane. as we can see in this picture.

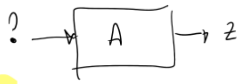
→ All the valid o/p's remain on the Im A (Plane)

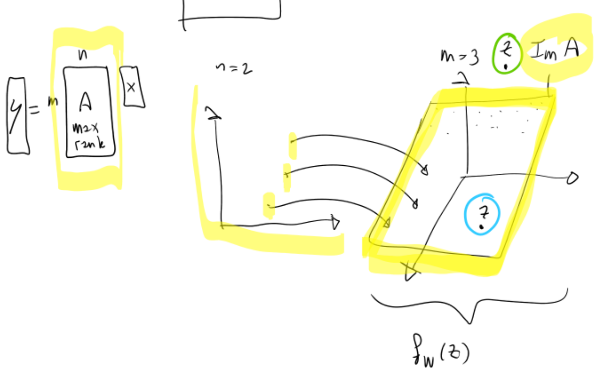
z cannot be produced ∴ this probability is Zero.

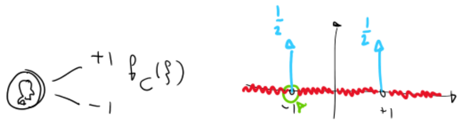
z have to follow the principle of Linear Transformation

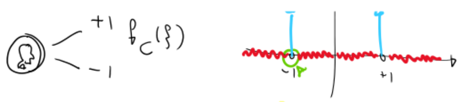
There is simply no problem in the no. of Counter Images since There is only one counter image for each point and since input 'x' is finite. Therefore we will have finite no. of counter Images

Non surjective $\exists z$ do not have a counterimage









$$f_c(\xi) = \frac{1}{2}\delta(\xi-1) + \frac{1}{2}\delta(\xi+\ )$$

$$f_w(z) = \delta(\ ) \boxed{}$$

checking whether the probability is zero

probability when not zero

$$f_w(z) = \delta((A^{\perp})^t z) \frac{f_x((A^tA)^{-1}A^t z)}{\sqrt{\det A^tA}}$$

orthogonal complement of A

n

Dear Shree
to your Amma's
Peddamma

జూలియా
जूलिया

జూలియా

जूलिया

जुलिया । । जुलिया

आलेख

Aa ? Alekh (

Aa

श्री

→ So, here in the case of Non-Surjective we know we have a finite Image i.e in the above example is a Plane Im A (all o/p's are on this plane)

Therefore we already have an idea how our o/p PDF is gonna Remain it's a PDF that is gonna remain on the plane and outside of it it should be zero probability.

On the plane o/p PDF has to obey the Rule layed out for the Domain & the Image which are both 2D-Planes. ∴ it is a Linear Transformation.

i.e it maps i/p plane onto o/p Image.

Q) How do we represent the o/p PDF which is present on a 2D plane in a 3D space?

We already know that o/p PDF is zero almost everywhere apart from the Plane (I on A)

We can make use of Delta function to eliminate the case of zero probability

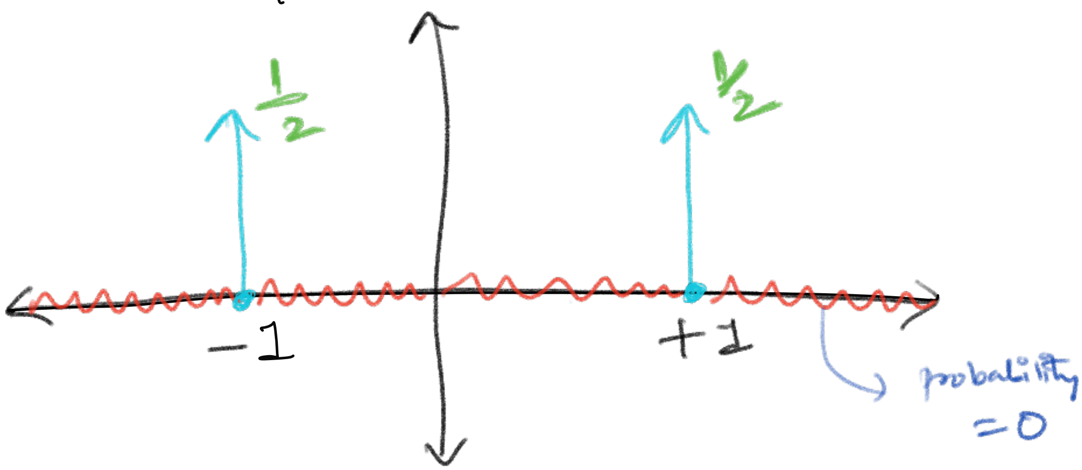
ex: Toin cossing example

Tails : +1

Heads : −1

To find the PDF of coin Tossing $f_c(q)$?

$q = +1$ (or) -1



→ PDF of the cointossing is zero every where except +1 , −1

$$\boxed{f_c(q) = \frac{1}{2}\delta(q-1) + \frac{1}{2}\delta(q+1)}$$

Non zero @ $\quad q = 1$

Non zero $\quad q = -1$

→ Therefore we can write the above structure with Delta functions to derive the o/p PDF of our system

The structure we can imagine is some thing like this

$$f_w(z) = \delta(\quad)\ \boxed{\quad}$$

Checking whether Probability = 0

probability when not zero

**

$$\boxed{f_w(z) = \frac{\delta\left(\left(A^{\perp}\right)^T z\right) f_x\left(\left(A^T A\right)^{-1} A^T z\right)}{\sqrt{\det(A^T A)}}}$$

① $A^{\perp} \rightarrow$ Orthogonality Complement of A ?

(it is a Matrix)

② Intresting Thing is figuring out what is The Orthogonality Compliment of A

$A^{\perp}$?

& what is $\underbrace{(AA^T)^{-1}}_{\text{This Matrix}} A^T z$?

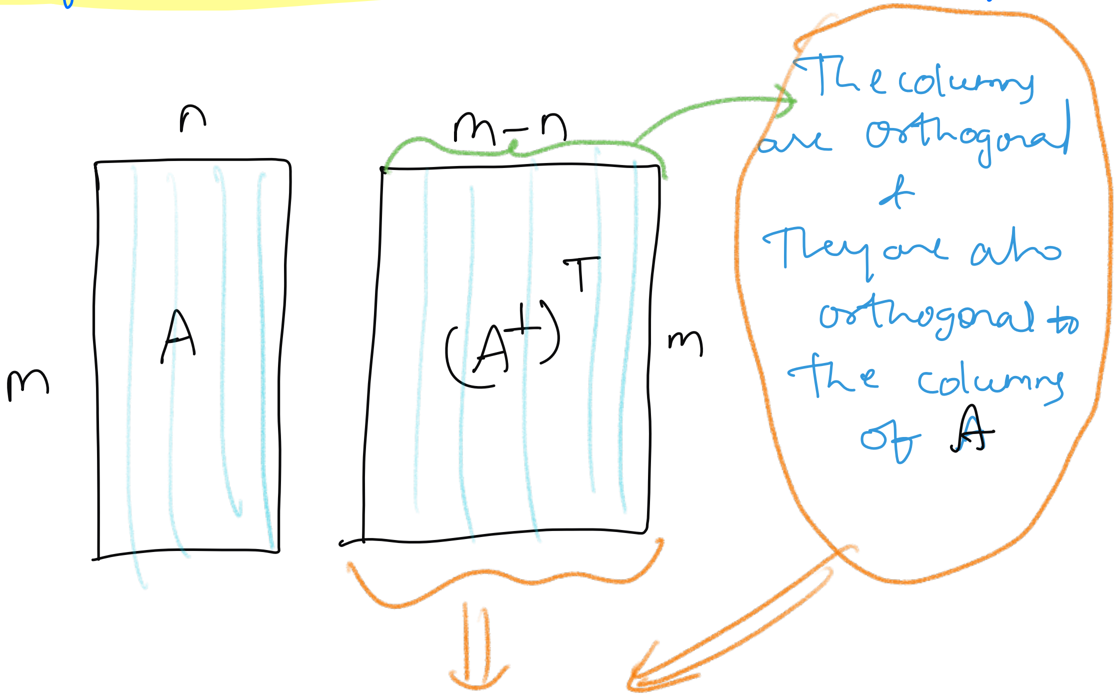
This the relative of pseudo Inverse we discussed earlier!

TALL MATRIX

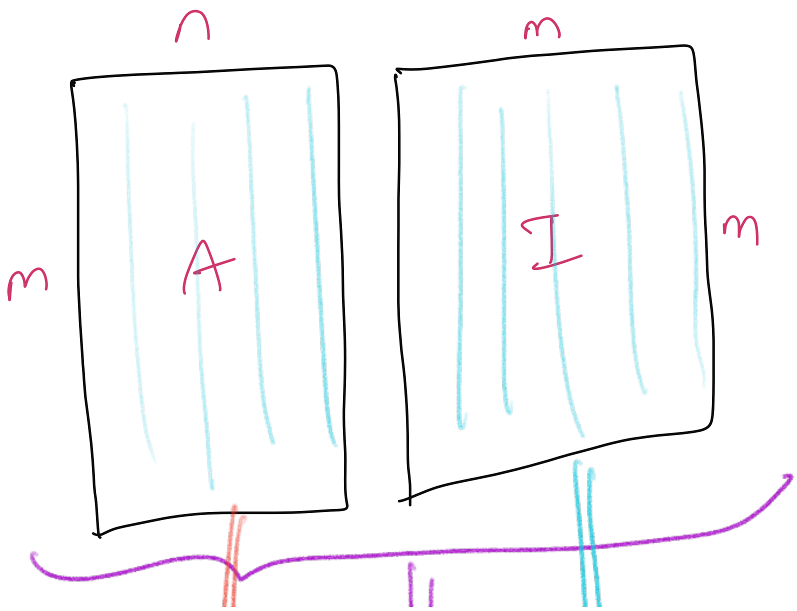
$m > n$ The no. of Columns is not enough for the basis of $\mathbb{R}^m$

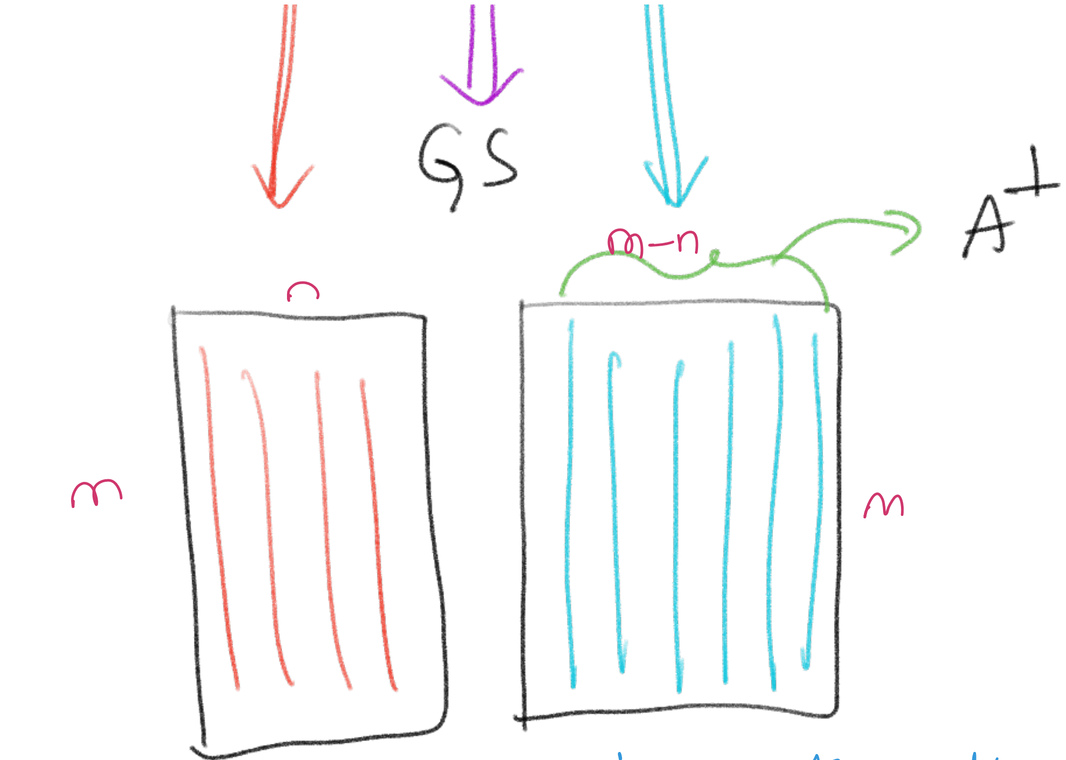
So, we have less columns than the dimensionality in which they live

** our Idea is to complete these columns with other columns, so that the union of columns of A & columns of $A^{\perp}$ s.t the columns of $A^{\perp}$ are orthogonal to the columns of A.



(columns of $A^{\perp}$) We can immediately realize that They can be produced by Gram-Schmidt Algorithm.





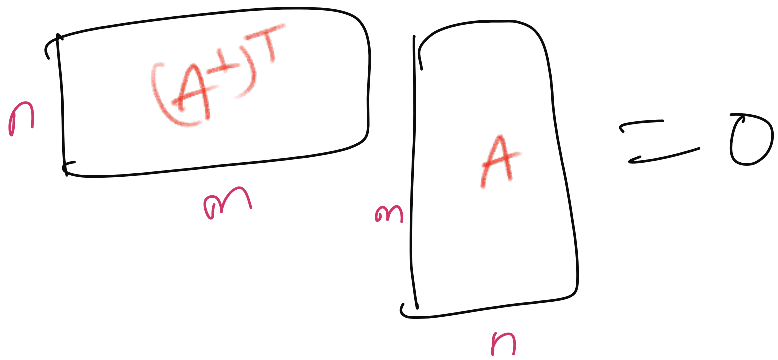
Basis of the colums of 'A'

columns orthogonal to the column of A

$\boxed{z \in \mathrm{Im}\, A}$ $\qquad z = Ay$

$$(A^{\perp})^T z = \underbrace{(A^{\perp})^T A}_{0}\, y = 0$$

it is non zero when we are out of Image of 'A'



Here is we do the scalar product of Rows of $(A^{\perp})^T$ & columns of A then the Result is Zero.

** ∴ it would be a perfect way to Test the if of the Delta function

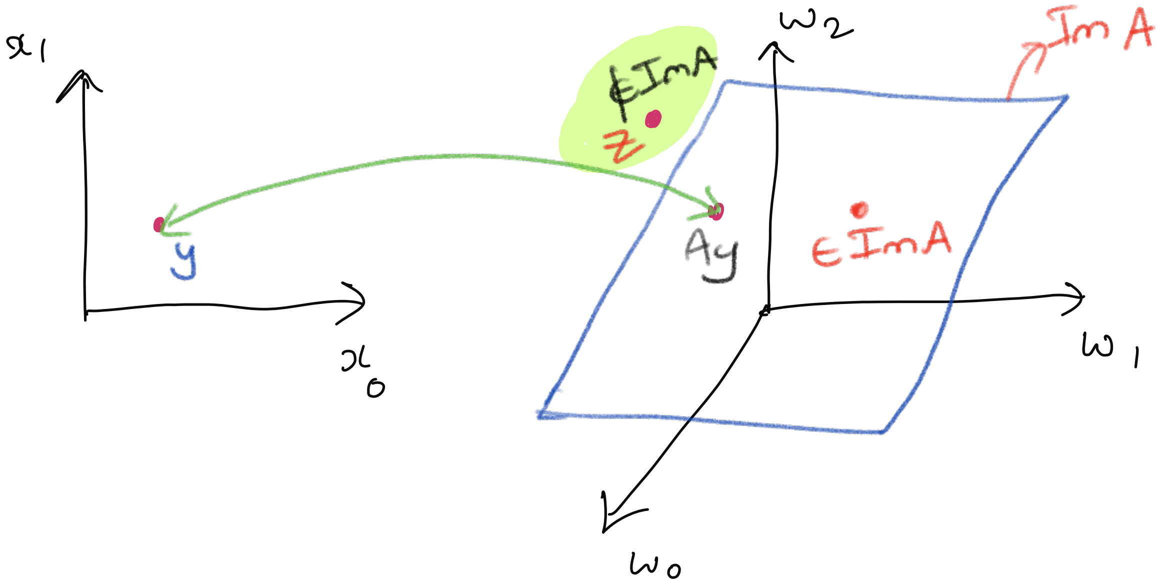
$$\text{i.e } f_w(z) = \delta\left((A^{\perp})^T z\right) \frac{f_X\left((A^TA)^{-1}A^T z\right)}{\sqrt{\det(A^TA)}}$$

→ Moore-Penrose Psuedo Inverse for Non-Surjective 'A'

- In the Non-Injective case we have Infinite no. of counter Images, so we needed a criteria to figure out the Infinite counter Images using the Ker A
- Similarly in this case of Non-Surjective we need to give counter images to the

points that does not have it

- Again, assuming the idea of distances we can have an example of mapping from $\mathbb{R}^2 \rightarrow \mathbb{R}^3$



- We can be sure that if this is a linear mapping, some of the points in $\mathbb{R}^3$ will be left out from the Im A & particularly 'z' as highlighted above does not belong to the Im A

- We have counter Image for every point on the Im A let us choose one point closest to 'z' that gives me least possible error! and assign the counterimage of that closest point as a CounterImage of 'z'. This is the closest approximation we can have in the Non-Surjective case!

We can solve for minimum possible error as an Optimization problem

Conjugate Transpose

$$\min \|Ay - z\|_2^2 = (Ay-z)^{+}(Ay-z)$$

Merit function is squared Norm of the vector

$$\Rightarrow (y^{+}A^{+} - z^{+})(Ay-z)$$

$$\Rightarrow y^{+}A^{+}Ay + z^{+}z \underbrace{- y^{+}A^{+}z - z^{+}Ay}$$

$$L(y) \Rightarrow \underbrace{y^{+}A^{+}Ay}_{\text{Quadratic form}} + z^{+}z - 2z^{+}Ay$$

This is the Lagrangian since we don't have to put any constraints

Taking Gradient w.r.t 'y'

$$\nabla_y L = 2A^{+}Ay - 2A^{+}z = 0$$

$$\Rightarrow A^{+}Ay = A^{+}z$$

$$y = (A^{+}A)^{-1}A^{+}z$$

expression for Counter Image of any point That is not in the Im A with least possible error!

This is pseudo inversion when 'A' is Non-Surjective

Note: This Inversion is possible because We assume Maximum Rank!

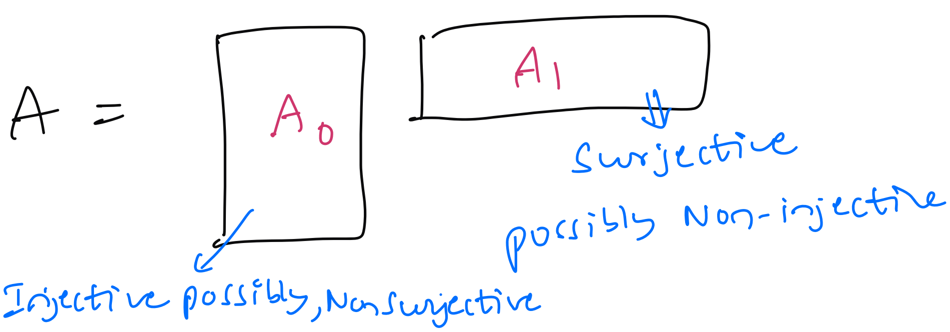
→ Non-Injective & Non Surjective

Till now we analyzed Non-injective & Non-Surjective case seperately, what we will do now is to Generalize the process for any Matrix (without assuming it as a Full Rank)

Any Matrix means ⟹ Noninjective + Non Surjective

BTW by addressing any possible Matrix, we will have a away of knowing of the PDF of the o/p that is produced by any Linear Algebraic processing.

This possibility relies on the property i.e Any matrix can be decomposed as product of Two Matrices



$$w = Ax = A_0 \underbrace{A_1 x}_{y}$$

$$\begin{tikzcd} x \arrow[r, "A_1"'] \arrow[rr, bend left, "A"] & y \arrow[r, "A_0"'] & w \end{tikzcd}$$

$$f_x \longrightarrow f_y \longrightarrow f_w$$

Non injective PDF formula — Non surjective PDF formula

→ If we can prove that A_0 & A_1 exist then we use above approach.

Finding A_0 & A_1

(I)

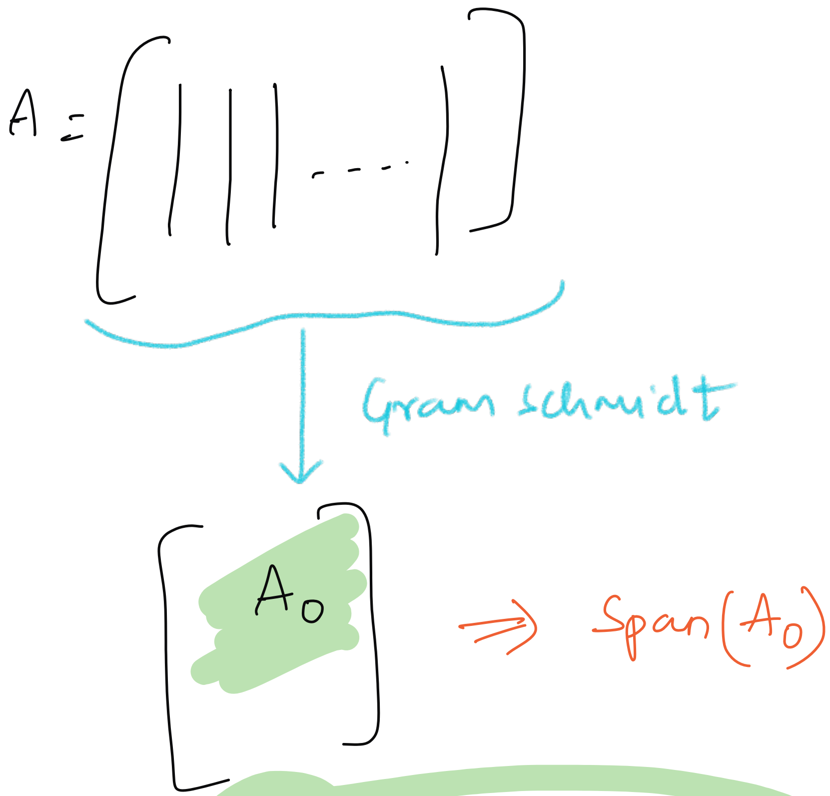
$$A = \begin{bmatrix} | & | & \cdots\cdots & | \\ | & | & & | \\ | & | & & | \end{bmatrix}$$

We know

$$\text{span}(A) = \text{Im}(A)$$

i.e linear combination of columns of 'A' we get the Image of 'A'

We can get the Basis of 'A' by using the Gram-schmidt i.e. If we feed the columns of 'A' into Gram-Schmidt then we get a Matrix that will contain Orthonormal Basis for our span



- We can say that $\text{Span}(A) = \text{Span}(A_0)$ because this is the property of Gram schmidt

This is the first step of our decomposition!

(II) 'A_0' is a potentially tall Maximum Rank matrix

We know how to make pseudoInverse for A_0:

Second step would be to show that A_1 can be found by multiplying 'A' with pseudoInverse of A_0

$$(A_0^t A_0)^{-1} A_0^t A$$ pseudo Inverse of A_0

$$\Rightarrow (A_0^t A_0)^{-1} A_0^t A_0 A_1$$

$$\Rightarrow \cancel{(A_0^t A_0)^{-1}} \cancel{(A_0^t A_0)} A_1$$

$$\Rightarrow A_1$$

Note: • This completes the statistical Analysis of Linear Algebraic process

Now we would like to do an example

ex: for an Nonjective case

We know that if we sum to Random Variables then the Pdf of the resultant R.V is the Convolution of the PDF of the RV.

$$f_w(z) = \frac{1}{\sqrt{\det(AA^t)}} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}} f_x\left[A^t (AA^t)^{-1} z + \sum_{j=0}^{n-m-1} \alpha_j e_j\right] d\alpha_0 \cdots d\alpha_{n-m-1}$$

$n-m$
integrals

Base vector of
ker(A)

$$\underset{\in \mathbb{R}^2}{x} \longrightarrow \boxed{\ A\ } \longrightarrow \underset{\in \mathbb{R}^1}{w}$$

$$\Rightarrow \quad w = Ax \quad \Rightarrow \quad w = \boxed{\quad A \quad}\, x$$

$$w = x_0 + x_1$$

$$w = \underbrace{\begin{bmatrix} 1 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

Now we want to find $f_w(z)$ using the Monster Formula for Non-Injective case, for that we need:

- Base vector of ker A
- pseudo inverse of A

(i) Kernel of A:

$$A = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

it's easy to find ker A by observation

$$\underbrace{\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{\ker A} = 0$$

$$e_0 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \Rightarrow$$ Scaled version of ker 'A' to make it orthonormal i.e length has to be ONE.

$$\| e_0 \| = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2} = 1$$

(ii) Psuedo Inversion of 'A'

$$A^t (AA^t)^{-1}$$

$$\therefore \quad AA^t = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2$$

$\det(AA^t) = 2$

$$A^t (AA^t)^{-1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{1}{2} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

$$\therefore f_\omega(z) = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} f_{x_0x_1}\left(\overset{\text{PsuedoInverse}}{\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}} z + \alpha_0 \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \right) d\alpha_0$$

$$f_{x_0x_1}\left(\begin{bmatrix} \frac{1}{2} z + \frac{1}{\sqrt{2}}\alpha_0 \\ \frac{1}{2} z - \frac{1}{\sqrt{2}}\alpha_0 \end{bmatrix} \right)$$

$$f_\omega(z) = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} f_{x_0x_1}\left(\frac{1}{2}z + \frac{1}{\sqrt{2}}\alpha_0 \,,\ \frac{1}{2}z - \frac{1}{\sqrt{2}}\alpha_0 \right) d\alpha_0$$

Here we rewrote the PDF as a function of two arguments

$$= \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} f_{x_0x_1}\left(\frac{1}{2}z + \frac{1}{\sqrt{2}}\alpha_0 \,,\ \underbrace{\frac{1}{2}z - \frac{1}{\sqrt{2}}\alpha_0}_{t} \right) d\alpha_0$$

$$t = \frac{1}{2}z - \frac{1}{\sqrt{2}}\alpha_0 \,, \quad dt = \frac{d\alpha_0}{\sqrt{2}}$$

Changing the variable of integration.

$$\boxed{f_w(z) = \int_{-\infty}^{\infty} f_{x_0 x_1}(\underbrace{z-t}_{x_0}, \underbrace{t}_{x_1})\, dt}$$

$$\underbrace{z-t}_{x_0} + \underbrace{t}_{x_1} = z$$

Here we are scanning for all values of 't' for which the sum of x_0, x_1 gives 'z'

- Now adding the assumption of x_0, x_1 being Independent

$$f_{x_0 x_1}(\eta_0, \eta_1) = f_{x_0}(\eta_0)\, f_{x_1}(\eta_1)$$

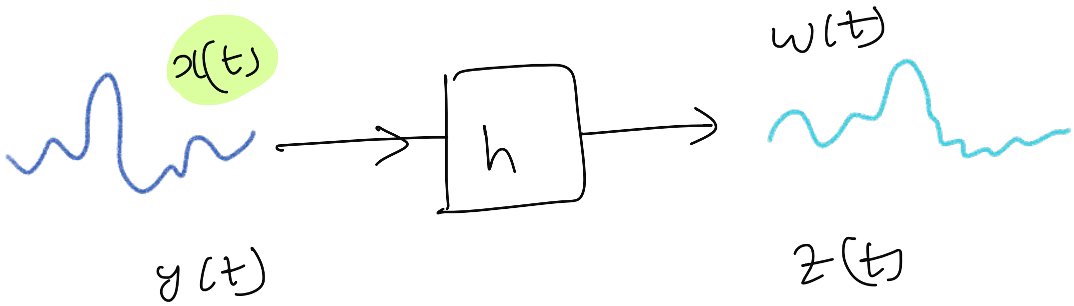
$$\boxed{f_w(z) = \int_{-\infty}^{\infty} f_{x_0}(z-t)\, f_{x_1}(t)\, dt}$$

Convolution of two PDFS

$$\boxed{f_w(z) = \left[f_{x_0} * f_{x_1}\right](z)}$$

Next step, Linear Dynamic Transformations

Goal: Is to find a way to characterize, How Stochastic signals are processed by linear means.



$x(t) \rightarrow$ input w.r.to time & we have also to consider the History of the process

$y(t) =$ Instance of the input

$w(t) \rightarrow$ o/p

$z(t) \rightarrow$ Instance of the o/p

→ The instances would be Waveforms & the Processes would be collection of waveforms with probability assigned to each waveforms.

→ The Idea that each waveform is mapped to another waveform in a linear way can be formalized by Impulse Response function

$$\boxed{z(t) = \int_{-\infty}^{\infty} \underline{h(t,s)}\, y(s)\, ds}$$

Impulse Response

Linear Mapping b/w i/p waveform to o/p waveform using

* We wanna know TWO things

1) How correlations are modified when processed through a Linear filter ?

2) How projections will be modied when processed through a linear filter ?

We should remember there are THREE ways of processing signals.

(i) Joint Probability ✗ (complicated)

(ii) Correlation & Covariance ✓

(iii) Projections. ✓

→ lets begin with Correlations & covariances would be exactly the same.

- The Idea is that we have to express the Statistical Feature of the o/p, based on the Statistical Feature of the Input & Characterization of the Block that make the processing. 'h'

The Statistical feature in this case is the Correlation Function

$$\underbrace{C_w(t_0, t_1, \ldots t_{m-1})}_{\text{Correlation function}} = E\left[z(t_0)\, z(t_1) \ldots z(t_{m-1})\right]$$

Of 'm' order which is a function of Timestamps

Note: We are dealing in Real Domain That is why we are not putting conjugates in the Expectation expression.

$$\Rightarrow \quad E\left[\int h(t_0, s_0)\, y(s_0)\, ds_0 \int h(t_1, s_1)\, y(s_1)\, ds_1 \ldots \ldots \int h(t_{m-1}, s_{m-1})\, y(s_{m-1})\, ds_{m-1}\right]$$

$$\Rightarrow \underbrace{\int \cdots \int}_{'m'} E\left[h(t_0,s_0)\,y(s_0)\,h(t_1,s_1)\,y(s_1)\cdots\cdots h(t_{m-1},s_{m-1})\,y(s_{m-1})\right] ds_0\,ds_1 \cdots ds_{m-1}$$

$\because$ Impulse Response is constant for a filter

$y(s)$, $z(t)$ are Random $\therefore$

$$\Rightarrow \boxed{\int \cdots \int h(t_0,s_0)\,h(t_1,s_1)\cdots h(t_{m-1},s_{m-1})\,\underbrace{E\left[y(s_0)\,y(s_1)\cdots\cdots y(s_{m-1})\right]}_{C_x(s_0,s_1,\ldots s_{m-1})} ds_0\,ds_1 \cdots ds_{m-1}}$$

it is nothing but Correlation function of the Input process.

* From the above Formula, we can say the Correlation of the o/p process can be expressed as Correlation of the i/p & the Response of the processing Box.

** Let us add TWO ASSUMPTIONS

(i) Linear Time Invariant Filters

$h(t,s) = h(t-s)$

(ii) Stationary Process

i.e

$x(t)$ is stationary

Time invariant in a statistical way is called Stationarity!

$$\forall T \quad \boxed{C_x(t_0, t_1, \ldots t_{m-1}) = C_x(t_0+T, t_1+T, \ldots\ldots t_{m-1}+T)}$$

$$C_w(t_0, t_1, \ldots t_{m-1}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \overbrace{h(t_0-s_0)\, h(t_1-s_1) \ldots \ldots h(t_{m-1}-s_{m-1})}^{LTI} C_x(s_0, s_1 \ldots s_{m-1})\, ds_0\, ds_1 \ldots ds_{m-1}$$

Changing the variables of Integration

let $t_J - s_J = q_J \Rightarrow dq_J = ds_J$

$s_J = t_J - q_J$

$$C_w(t_0, t_1 \dots t_{m-1}) \Rightarrow \int \dots \int h(\tau_0) h(\tau_1) \dots h(\tau_{m-1}) \, C_x[t_0-\tau_0, t_1-\tau_1, \dots t_{m-1}-\tau_{m-1}] \, d\tau_0 \, d\tau_1 \dots d\tau_{m-1}$$

If choose the Temporal shift $T = -t_0 + \tau_0$

$$C_x(t_0-\tau_0, t_1-\tau_1, \dots t_{m-1}-\tau_{m-1})$$

$$= C_x(t_0-\tau_0+T, t_1-\tau_1+T, \dots, t_{m-1}-\tau_{m-1}+T)$$

$$\Downarrow$$

$$\Rightarrow C_x(0, t_1-t_0+\tau_0-\tau_1, \dots, t_{m-1}-t_0+\tau_0-\tau_{m-1})$$

$$\Rightarrow C_w(t_0, t_1 \dots, t_{m-1})$$

$$= \int \dots \int h(\tau_0) h(\tau_1) \dots h(\tau_{m-1}) \quad t_J - t_0 + \tau_0 - \tau_J$$

$$C_x(0, t_1-t_0+\tau_0-\tau_1, \dots \dots t_{m-1}-t_0+\tau_0-\tau_{m-1})$$

$$d\tau_0 \, d\tau_1 \cdots\cdots d\tau_{m-1}$$

Now, we want to check whether correlation o/p is stationary or not.

$$C_w(t_0+T, t_1+T, \ldots t_{m-1}+T)$$

$$= C_w(t_0, t_1 \ldots t_{m-1})$$

Because of $t_J - t_0 + \tau_0 + \tau_J$

$$\Downarrow$$

$$t_J + \cancel{T} - t_0 - \cancel{T}$$

The whole expression would remain same

This is generally what we like for the process to behave. i.e when we take a stationary process and feed it into a LTI filter what comes out is a stationary process.

Note: The problem is it is NOT TRUE in General

Ex:

For coin tossing example

x_t (R.V) is ± 1

$$\left.\begin{array}{l} \text{Head} \longrightarrow +1 \\ \text{Tails} \longrightarrow -1 \end{array}\right\} \text{Stationary process}$$

• Now let's take a simple Filter (INTEGRATOR)

$$w_t = w_{t-1} + x_t$$

Integrator is Time Invariant

current o/p — w_t

previous o/p — w_{t-1}

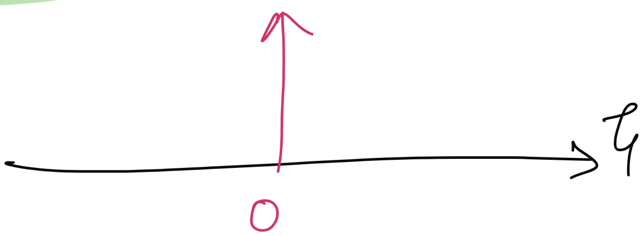
current input — x_t

For suppose, at $t=0$ $w_t = 0$

PDF @ $t=0$

$$\boxed{f_{w(0)}(\eta) = \delta(\eta)}$$

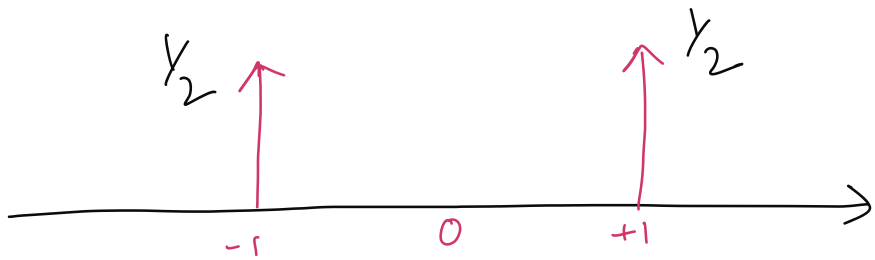
we know at $\eta = 0$, the whole probability is concentrated. and the PDF is shown below.



• @ $t=1$ $w_1 = w_0 + x_1$ → +1 Heads, −1 Tails

PDF @ $t=1$

$$\therefore \boxed{f_{w(1)}(\eta) = \frac{1}{2}\delta(\eta+1) + \frac{1}{2}\delta(\eta-1)}$$



We can see that the PDF @ $t=0 \neq$ PDF @ $t=0$

$\therefore$ This process is not stationary

It's Integrator's Fault!

→ If we don't put constraints on Impulse Response we can guarantee the Theorem which interchanges Expectation with an Integration.

Bertlokki Move Because we may have convergence issues!

* Second order Correlations ($m=2$)

$$C_{w}(t_0, t_1) = \iint h_0(t_0, s_0)\, h(t_1, s_1)\, C_x(s_0, s_1)\, ds_0\, ds_1$$

it is Stationary & LTI