X_N S_N

In ICA problem is to obtain $\underline{S}$, Given $\underline{X}$

$$\underline{S} = A^{-1}\underline{X} \qquad \text{let } A^{-1} = W$$

$$\therefore \underline{S} = W\underline{X}$$

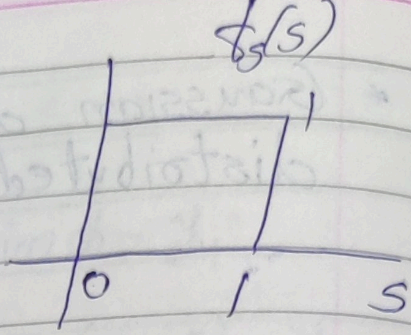
- $\underline{S}$ itself is a linear combination of Given $\underline{X}$.

* How to get $\underline{W}$ (Background):-

- Consider a random variable S and obtain $X = aS$, a is a constant, then

$$f_X(x) = \frac{f_S(s)}{|a|} = \frac{f_S\left(\frac{x}{a}\right)}{|a|}$$

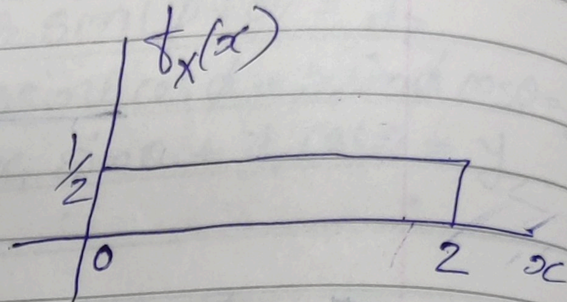
e.g. $x = 2s$



$\rightarrow$ then $f_x(x) = \frac{f_s(s)}{2} = \frac{f_s(\frac{x}{2})}{2}$

$0 \le s \le 1$

$\therefore 0 \le x \le 2$



- If s_1, s_2 are RVs. and x_1 and x_2 are linear transformation of s_1 and s_2 then

$$x_1 = a_{11}s_1 + a_{12}s_2$$
$$x_2 = a_{21}s_1 + a_{22}s_2$$

$$f_{x_1,x_2}(x_1, x_2) = \frac{f_{s_1,s_2}(s_1, s_2)}{|A|}$$

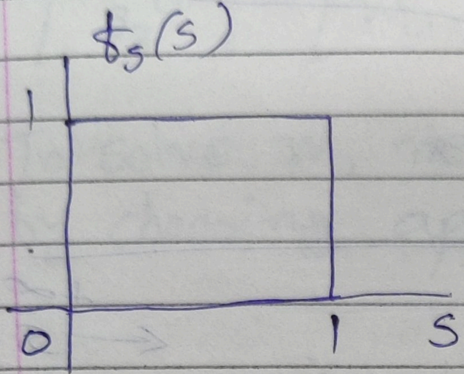
$\hookrightarrow$ Determintant of matrix and mod of that

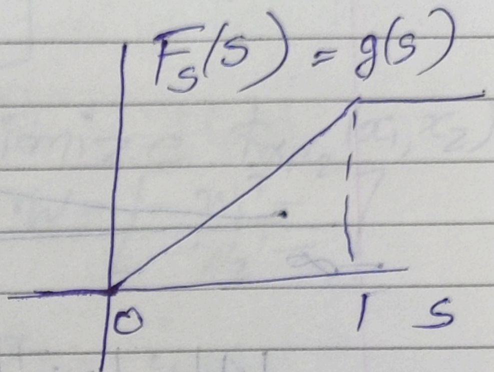
$\rightarrow$ If RVs $s_1, s_2, \dots, s_N$ are statistically independent then

$$f_{s_1 s_2 \dots s_N}(s_1, s_2, \dots, s_N) = f_{s_1}(s_1)\, f_{s_2}(s_2) \dots f_{s_N}(s_N)$$

- Assume a density f^n for $s_1, s_2, \dots, s_N$ vector RV with multivariate density function.

* Cummulative distribution ($F_s(s)$):-





→ $f_s(s) = g'(s) = \frac{d}{ds}(F_s(s))$

$F_s(s_1) = P(S \leq s_1)$

$= \int_{-\infty}^{s_1} f_s(t)\, dt$

→ We know that $\underline{S} = W\underline{X}$ where $W = A^{-1}$

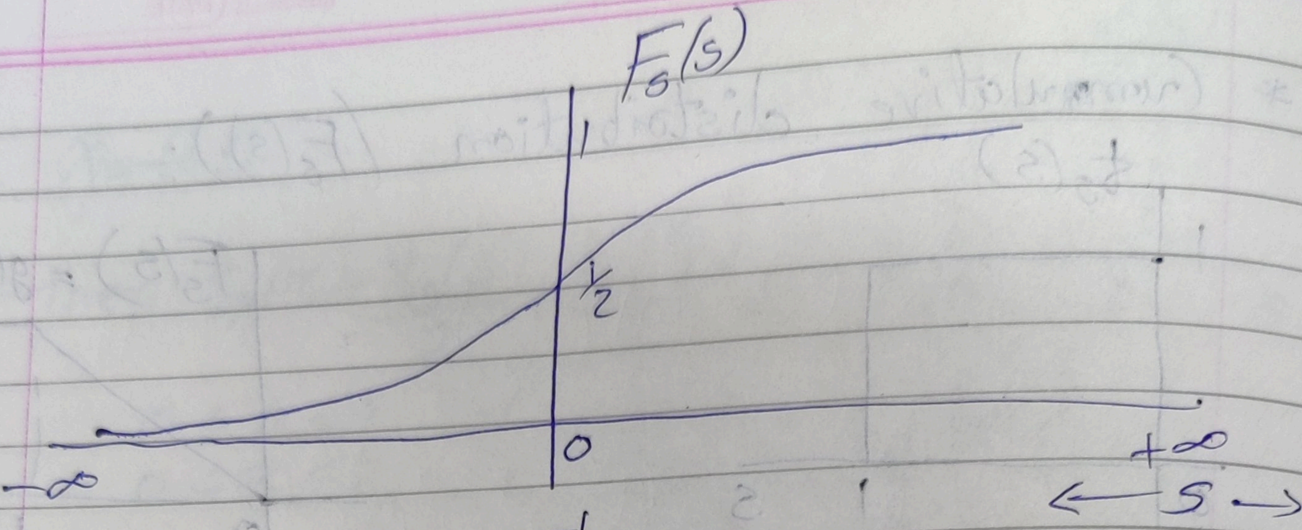
$\underline{X} = A\underline{S}$

- For $N = 2$

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

- let us assume $g(s) = F_s(s) = \frac{1}{1+e^{-s}}$



$$|A| = |w^{-1}| = |w|^{-1}$$

- From definition of pdf of multivariate density f^n

$$f_{x_1 x_2 \dots x_N}(x_1, x_2, \dots, x_N) = f_{s_1 s_2 \dots s_N}(s_1, s_2, \dots, s_N) \cdot |w|$$

- if $N = 2$

$$f_{x_1 x_2}(x_1, x_2) = f_{s_1 s_2}(s_1, s_2)\ |w|$$

$$= f_{s_1}(s_1)\, f_{s_2}(s_2)\ |w|$$

- $W = \begin{bmatrix} -w_1^T- \\ -w_2^T- \\ \vdots \\ -w_N^T- \end{bmatrix}$ $\quad w_1 = \begin{bmatrix} w_{11} \\ w_{12} \end{bmatrix}$

$$\therefore f_{x_1 x_2}(x_1, x_2) = f_{s_1}(w_1^T x)\, f_{s_2}(w_2^T x)\ |w|$$

$\rightarrow$ $S_1 = w_{11} x_1 + w_{12} x_2$
$S_2 = w_{21} x_1 + w_{22} x_2$

$X_1 = x_1^{(1)}\ x_1^{(2)} \dots x_1^{(m)}$
$X_2 = x_2^{(1)}\ x_2^{(2)} \dots x_2^{(m)}$

$$\begin{bmatrix} s_1^{(1)} \\ s_2^{(1)} \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix}$$

- To solve w, we will maximize $f_{x_1 x_2}(x_1, x_2)$ by choosing appropriate $w = \begin{bmatrix} - w_1^T - \\ - w_2^T - \end{bmatrix}$

- $f_{x_1 x_2}(x_1, x_2) = f_{s_1}(w_1^T x) \, f_{s_2}(w_2^T x) \, |w|$

$f_{s_1}(w_1^T x) = g'(w_1^T x)$

$f_{s_2}(w_2^T x) = g'(w_2^T x)$

~~x_1, x_2 and~~

- We have to assume some density function of s. and we can write the log likelihood f^n as

$$\max_{\forall w} \; l(w) = \sum_{i=1}^{m} \left(\sum_{j=1}^{N} \log\left(g'(w_j^T x^{(i)})\right) + \log|w| \right)$$

→ We can maximize $l(w)$ or minimize $-l(w)$ by using stocastic gradient descent

→ For stocastic gradient descent

$$l(w) = -\left[\sum_{j=1}^{N} \log\left(g'(w_j^T x^{(i)})\right) + \log|w| \right]$$

PAGE NO.
DATE

→ let us consider differentiation w.r.t. $w_{11}, w_{12}, w_{21}, w_{22}$ where $n = 2$

$$\therefore l(W) = \sum_{j=1}^{2} \log\left(g'(w_j^T x^{(i)})\right) + \log|W|$$

$$= \log\left(g'(w_1^T x^{(i)})\right) + \log\left(g'(w_2^T x^{(i)})\right) + \log|W|$$

$$= \log\left(g(w_1^T x^{(i)})\left(1 - g(w_1^T x^{(i)})\right)\right)$$

$$+ \log\left(g(w_2^T x^{(i)})\left(1 - g(w_2^T x^{(i)})\right)\right) + \log|W|$$

$$\rightarrow \frac{\partial l(W)}{\partial w_{11}} = \left(1 - 2g(w_1^T x^{(i)})\right) \cdot x_1^{(i)}$$