PASFML, Mansoura AI45 11/1/2025

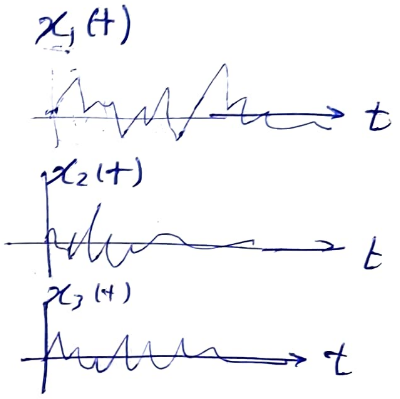
Session 6

→ Markov process / Markov chain

→ Entropy (Information Entropy)

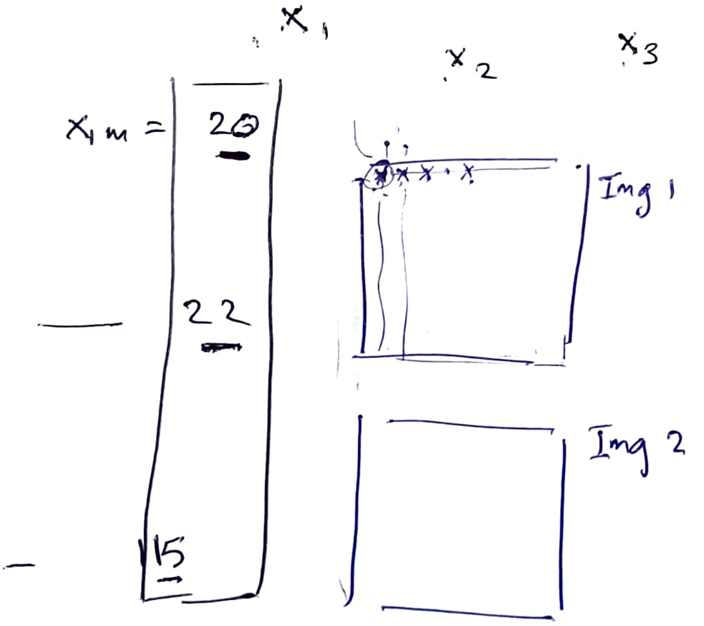
Random Process (stochastic Process)

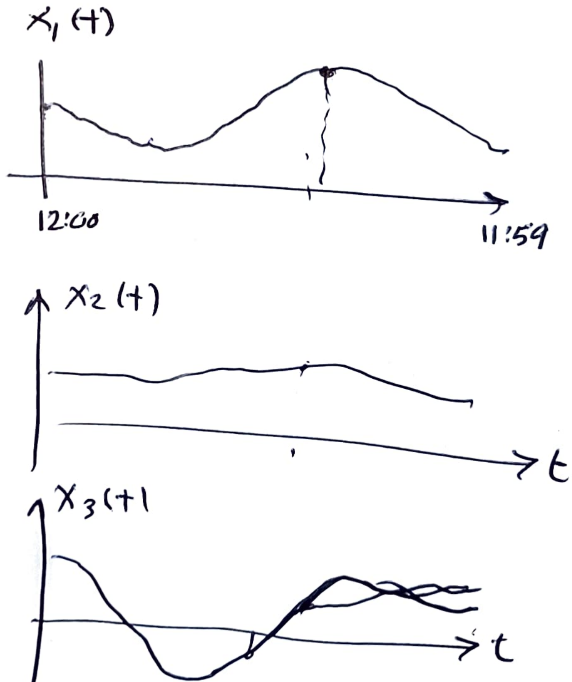
v/s (vs. Random variable) $\Leftarrow X = x_1$ ✓



$X(\xi; t)$

ξ ↓ Random variable





Images : Random process

"spatial domain"

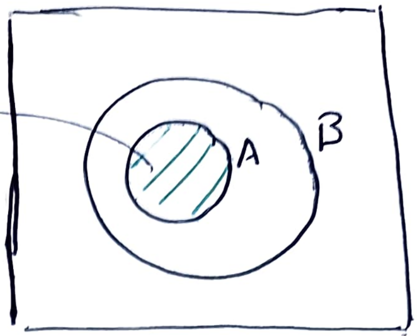
space

ensemble of ~~different~~ possible Realizations of process

"Review"

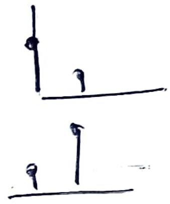
$P(A) = 1/4$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



$P_X(1) = p \qquad P_X(0) = 1 - p = q$

$P(X=1) = p = 1 - P(X=0) = 1 - q$

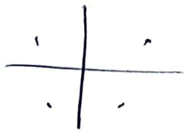


$E[X] = \sum_i x\, P(x)$

$= 0 \times P_X(0) + 1 \times P_X(1) = p = 1 - q$

$X = \cos\Phi \qquad Y = \sin\Phi \rightarrow$ not indep.

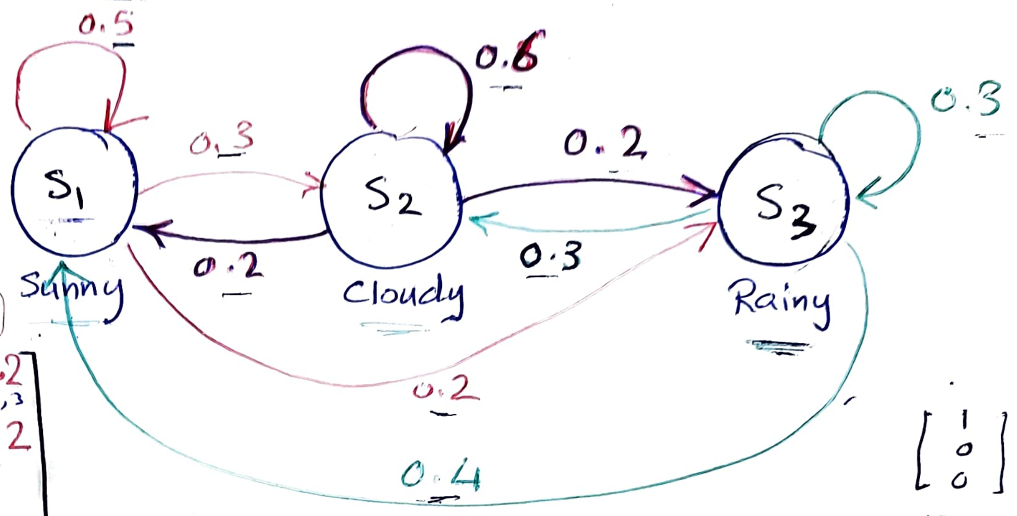
$Y = \sin(\cos^{-1} X)$



Discrete (time) Markov Random Process:

memoryless ✓

short-term memory



$$
\begin{array}{c|c}
 & \begin{matrix} S_1^+ & S_2^+ & S_3^+ \end{matrix} \\
\begin{matrix} S_1 \\ S_2 \\ S_3 \end{matrix} & \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.4 & 0.3 & 0.3 \end{bmatrix}
\end{array}
$$

next state: S_1^+, S_2^+, S_3^+

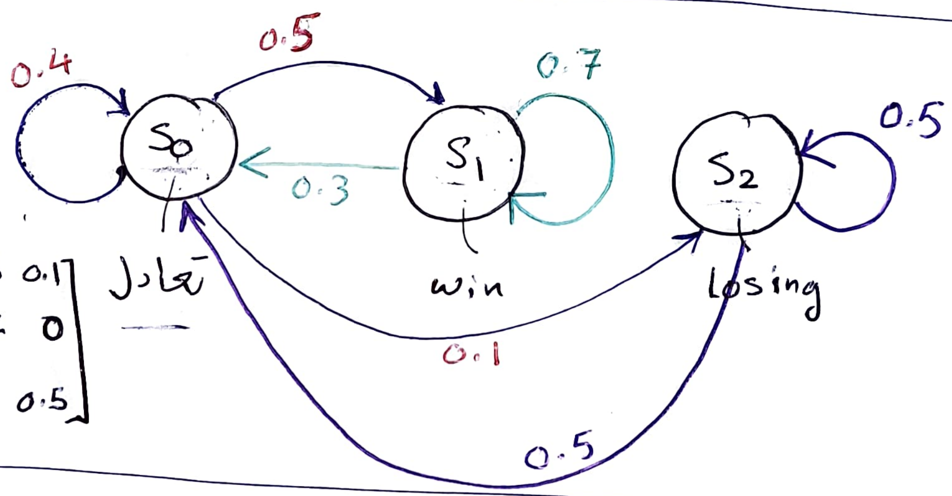
$P_{1,1}$ $P_{1,2}$ $P_{1,3}$

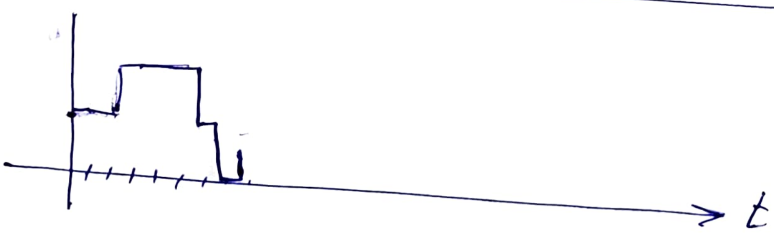
current state: S_1, S_2, S_3

← transition Probability matrix

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ $\begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix}$

$$\begin{bmatrix} 0.4 & 0.5 & 0.1 \\ 0.3 & 0.7 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$



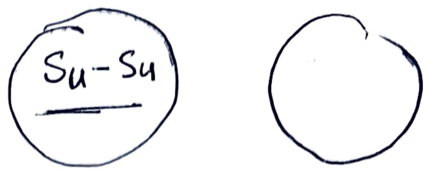


$X^{(0)}$ $X^{(1)}$

next state vector $\vec{X}^+ = \underline{A}\,\vec{X}$

initial state vector $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ ← Initially

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}}_{\text{current state}} \begin{bmatrix} 0.4 & 0.5 & 0.1 \\ 0.3 & 0.7 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix} = \underbrace{\begin{bmatrix} 0.4 & 0.5 & 0.1 \end{bmatrix}}_{\text{next state vector}}$$

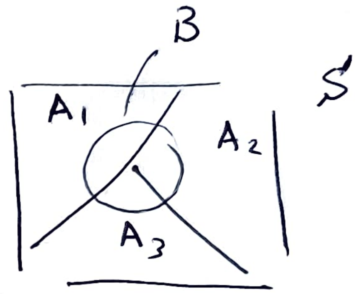


Recall conditional Probability.

$$P(B/A) = P(BA)/P(A)$$

$$P(BA) = P(B/A)\,P(A)$$

$$P(B) = P(BA_1) + P(BA_2) + P(BA_3)$$



$P_{A_1,B}$

$$\boxed{\begin{aligned} P(B) = {} & P(B/A_1)\,P(A_1) \\ & + P(B/A_2)\,P(A_2) \\ & + P(B/A_3)\,P(A_3) \end{aligned}}$$

$$P(B) = \sum_i P(B/A_i)\,P(A_i)$$

$$P(S_j^+) = \boxed{\sum_i P(S_j^+/S_i)\,P(S_i)}$$

$S_1^{(0)}$

$S_1^{(1)}$

$S_1^{(2)}$

$\vdots$

$P_{0,0} \equiv P(S_0^+/S_0)$

$P_{0,1} \equiv P(S_1^+|S_0)$

$P_{i,j} \equiv P(S_j^+/S_i)$

notations of Markov transition Probabilities.

$$[P_0 \quad]\begin{bmatrix} P(S_1) \\ P(S_2) \\ \vdots \end{bmatrix}$$

$$\begin{aligned} & P(S_j^+/S_0)\,P(S_0) \\ & + P(S_j^+/S_1)\,P(S_1) \\ & + P(S_j^+/S_2)\,P(S_2) \end{aligned}$$

$$\equiv P_{0,j}\,P(S_0) + P_{1,j}\,P(S_1) + P_{2,j}\,P(S_2) = P(S_j^+)$$

$$[\quad]\,[\quad]$$

$$P = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} \\ P_{1,0} & P_{1,1} & P_{1,2} \\ P_{2,0} & P_{2,1} & P_{2,2} \end{bmatrix}$$

$$\cancel{P(0)^{+} = P_{0,0} \cdot P(0) + P_{1,0} \cdot P(1) + P_{0,2}}$$

$$\boxed{P(S_0^{+}) = P_{0,0} \times P(S_0) + P_{1,0} \times P(S_1) + P_{2,0}\ P(S_2)}$$

$$\vec{X}^{+} = \boxed{\begin{bmatrix} P(S_0) & P(S_1) & P(S_2) \end{bmatrix}} \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} \\ P_{1,0} & P_{1,1} & P_{1,2} \\ P_{2,0} & P_{2,1} & P_{2,2} \end{bmatrix}$$

$$\begin{bmatrix} P(S_0^{+}) & P(S_1^{+}) & P(S_2^{+}) \end{bmatrix} =$$

next state "column number"

(5)

$$P = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.4 & 0.3 & 0.3 \end{bmatrix}$$

current state "row number"

$$P = [\leftarrow P_{i,j} \rightarrow]$$

Initially

$$\vec{X}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\vec{X}^{(1)^T} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.4 & 0.3 & 0.3 \end{bmatrix}$$

$$\vec{X}^{(1)^T} = \begin{bmatrix} 0.5 & 0.3 & 0.2 \end{bmatrix}$$

$$\vec{X}^2 = \vec{X}^{(1)^T} [P] = [0.5 \quad 0.3 \quad 0.2] \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

$$[x \quad x \quad x]$$

6

Current state index = column number

IF WE TRANSPOSED THE PROPABILITY MATRIX TO BE:

$$P = \begin{bmatrix} P_{0,0} & P_{1,0} & P_{2,0} \\ P_{0,1} & P_{1,1} & P_{2,1} \\ P_{0,2} & P_{1,2} & P_{2,2} \end{bmatrix}$$

next state index = row number

$$\vec{x}^{(1)} = \begin{bmatrix} \\ \\ \end{bmatrix} = \begin{bmatrix} & & \\ & & \end{bmatrix} \underbrace{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}}_{\vec{x}^{(0)}}$$

$$\vec{x}^{(n)} = P^n \vec{x}^{(0)}$$

$$P \times P \times \dots \times P \leftarrow n \text{ times.}$$

if P is diagonalizable.

$$P \times P = [\vec{v}] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_i \end{bmatrix} \underbrace{[\vec{v}]^{-1} [\vec{v}]}_{I} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_i \end{bmatrix} [\vec{v}]^{-1}$$

$$P^n \vec{x}^{(0)} = [\vec{v}] \begin{bmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_i^n \end{bmatrix} [\vec{v}]^{-1} \vec{x}^{(0)} = \vec{x}^{(n)}$$

---

steady state? هل الاحتمالات تثبت على المدى الطويل

$$\vec{x}^{(n)} = \underbrace{\begin{bmatrix} & \\ & \end{bmatrix}}_{P} \vec{x}^{(n-1)} = \vec{x}^{(n-1)}$$

$$P \vec{x} = 1 \vec{x}$$

eigendecomposition

$$\lambda_i = 1 \Rightarrow \vec{v}_i$$

7

$$
\begin{tikzcd}
S_0 \arrow[r, "1"] & S_1 \arrow[loop above, "0.25"] \arrow[r, "0.75"] & S_2 \arrow[loop above, "0.5"] \arrow[r, "0.5"] & S_3 \arrow[loop above, "0.75"] \arrow[r, "0.25"'] & S_4
\end{tikzcd}
$$

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

$$
\vec{x}^{(6)} = P^6 \, \vec{x}^{(0)} \qquad \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
$$

Information Entropy

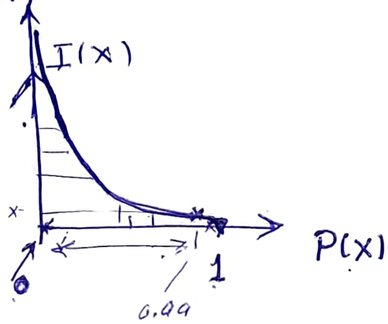
01010100111010110 $\qquad$ 0000010000000101

$$I(x) \propto \frac{1}{f(P(x))}$$

$I(x)$ total, (x_1, x_2) $= I(x_1) + I(x_2)$

$$I(x) = \log_2 \frac{1}{P(x)} \text{ bits} = -\log P(x)$$

Information



Information entropy.

$P(H) = 0.5$

$P(T) = 0.5$

$I(H) = -\log_2 (0.5)$

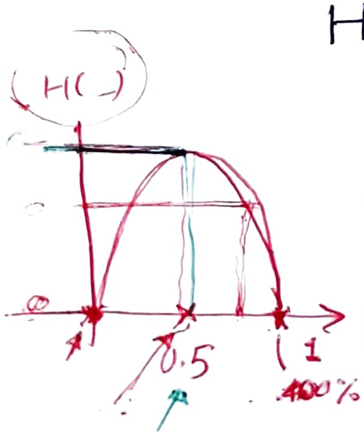
$I(T) = -\log_2 (0.5)$

average information $= \sum_i I(x_i) P(x_i)$

Entropy

$$H(x) = \sum_i (-\log_2 P(x_i)) P(x_i)$$

$$H(x) = -\sum_i -P(x_i) \log P(x_i)$$



$0.7 \qquad\qquad 0.3$

$-P(H) \log P(H) - P(T) \log(T)$

Cross-Entropy loss function (Classification)

$$H(P,q) = -\sum_i P(x_i) \log q(x_i)$$

true Prob.

prediction prob.

1

1

0.6

0.4