Stochastic Process: Lect 6 & 7

Intro to Markov Chain.

What is stochastic process?

A family of Random variables X_n that is indexed by $\underline{\underline{n}}$. $n \in T$.

① T is discrete: $T = \{0, 1, 2, 3 \cdots\}$

discrete time stochastic process.

② T is continuous: $T = [0, +\infty)$

continuous time stochastic process.

X_n can have a discrete / continuous distribution.

usually X_n takes its value in the same set.

$X_n \in S$

state space

Let $\{X_n : n = 0, 1, 2 \cdots\}$ be a discrete time stochastic process and state space S is discrete.

The model is describible by the joint distribution

$$P(X_0 = \underbrace{x_0}_{\in S}, X_1 = \underbrace{x_1}_{\in S}, \cdots X_n = \underbrace{x_n}_{\in S})$$

In i.i.d case $= \prod_{i=0}^{n} P(X_i = x_i)$

① Excess # of heads over tails in tossing a coin. (assume each coin is tossed independently).

X_n $\quad S$

$X_n \in \{ \cdots -2, -1, 0, 1, 2, \cdots \}$

Then the process $\{X_n\}$ has one-step memory.

$$P(X_0, \cdots X_n) = P(X_n | X_{n-1}) P(X_{n-1} | X_{n-2}) \cdots P(X_0)$$

It is not in i.i.d.

Markov chain. (current state X_{n-1} captures all relevent information about the state)

A dt stochastic process is a Markov Chain

if for every $x_0, x_1, \cdots, x_{n-2}, i, j \in S, n > 0$

$$P(X_n = i | X_{n-1} = j, \cdots X_0 = x_0)$$

$$= P(X_n = i | X_{n-1} = j)$$ one step memory

called Markov assumption

if $\triangleq P_{ji}$ (is independent of n)

Then X_n is a $\boxed{\text{homogeneous Markov Chain}}$.

property: $\sum_{i \in S} P(X_n = i \mid X_{n-1} = j) = 1.$

$\Leftrightarrow \sum_{i \in S} P_{ji} = 1$ for all j.

and $P_{ji} \geq 0$

prob moving from state j to state i.

Label it as $S = \{1, 2, \cdots, N\}$

the element $\{P_{ji} : j, i = 1, 2, \cdots, N\}$

forms an $N \times N$ Matrix $P = \{P_{ij}\}$

$\hookrightarrow$ called transition prob. Matrix

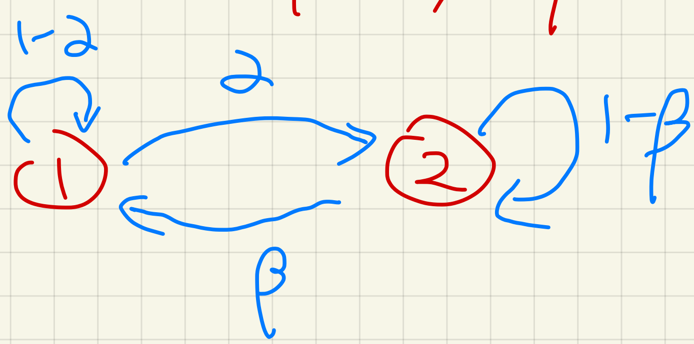
property: $P \geq 0$ and $\mathbb{1}^T P = \mathbb{1}^T$ (stochastic Matrix).

One can visualize such Markov chain by drawing a directed graph.

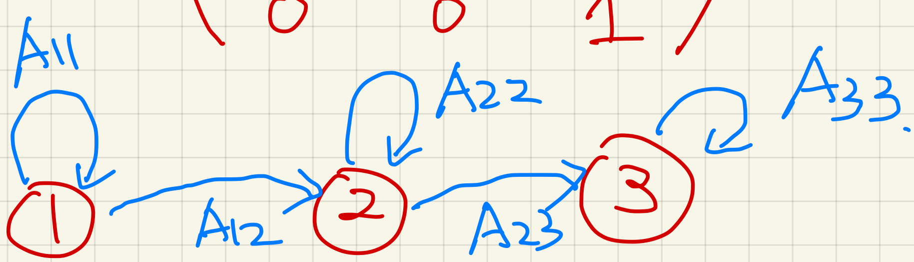
node: states

arrow: legal transitions (non zero elements in P)

Ex: $P = \begin{pmatrix} 1-\alpha, & \alpha \\ \beta, & 1-\beta \end{pmatrix}$



$P = \begin{pmatrix} A_{11} & A_{12} & 0 \\ 0 & A_{22} & A_{23} \\ 0 & 0 & 1 \end{pmatrix}$ left-to-right transition Matrix



Markov Property: distribution of X_{m+n} (Markov chain) given a set of previous states depends on the latest available state, i.e.

Thm 1: $P(X_{m+n}=j \mid X_n=i, X_{n-1}=i_{n-1}, \cdots X_0=i_0)$

$= P(X_{m+n}=j \mid X_n=i)$.

we prove the special simple case when $n=1$ and $m=2$.

$$P(X_3=i_3 \mid X_1=i_1, X_0=i_0)$$

$$= \sum_{i_2} P(X_3=i_3, X_2=i_2 \mid X_1=i_1, X_0=i_0)$$

$$= \sum_{i_2} P(X_3=i_3 \mid X_2=i_2, X_1=i_1, X_0=i_0) \cdot P(X_2=i_2 \mid X_1=i_1, X_0=i_0).$$

$$= \sum_{i_2} P(X_3=i_3 \mid X_2=i_2, X_1=i_1) P(X_2=i_2 \mid X_1=i_1).$$

$$= \sum_{i_2} P(X_3=i_3, X_2=i_2 \mid X_1=i_1)$$

$$= P(X_3=i_3 \mid X_1=i_1)$$

$$0 - 1 - 2 \cdots\cdots - n-1 - n$$

$n=2$, X_0: past

X_1: present

X_2: future

Thm 2: X_0 and X_2 are conditional independent given X_1. for any $i_0, i_1, i_2 \in S$

i.e. $P(X_0 = i_0, X_2 = i_2 \mid X_1 = i_1)$

$$= P(X_0 = i_0 \mid X_1 = i_1)\, P(X_2 = i_2 \mid X_1 = i_1)$$

proof: $P(X_0 = i_0, X_2 = i_2 \mid X_1 = i_1)$

$$= P(X_0 = i_0, X_1 = i_1, X_2 = i_2) / P(X_1 = i_1)$$

$$= \frac{P(X_2 = i_2 \mid X_1 = i_1, X_0 = i_0)\, P(X_1 = i_1, X_0 = i_0)}{P(X_1 = i_1)}$$

$$= P(X_2 = i_2 \mid X_1 = i_1)\, P(X_0 = i_0 \mid X_1 = i_1)$$

n-step transition probability for MC.

is $P_{ij}^{(n)} = P(X_n = j \mid X_0 = i)$

Lemma: Let $\{X_n\}$ be a homogeneous Markov chain and $P_{ij}^{(n)}$ be n-step transition prob. then for any $k = 0, 1, 2, \cdots$

$$P(X_{n+k} = j \mid X_k = i) = P_{ij}^{(n)}$$

proof:

$$P(X_{n+k} = j \mid X_k = i) = \sum_{i_{k+1}, i_{k+2}, \cdots i_{n+k-1}} P(X_{n+k} = j \mid X_{n+k-1} = i_{n+k-1})$$

$$\times \cdots \times P(X_{k+1} = i_{k+1} \mid X_k = i)$$

homogeneous property

$$= \sum_{i_{k+1}, \cdots i_{n+k-1}} P(X_n = j \mid X_{n-1} = i_{n+k-1}) \times \cdots \times P(X_1 = i_{k+1} \mid X_0 = i)$$

$$= P(X_n = j \mid X_0 = i) = P_{ij}^{(n)}$$

n-step transition prob are related to each other via Chapman-Kolmogorov equation.

Lemma:

Let $\{X_n\}$ be a homogeneous MC and Let $P_{ij}^{(n)}$ be the n-step transition probability.

Then for any $n, m = 0, 1, 2, \cdots$

$$P_{ij}^{(n+m)} = \sum_{k \in S} P_{ik}^{(n)} P_{kj}^{(m)}$$

proof:

$$P_{ij}^{(n+m)} = P(X_{n+m} = j \mid X_0 = i)$$

$$= \sum_{k \in S} P(X_{n+m} = j, X_n = k \mid X_0 = i)$$

$$= \sum_{k \in S} P(X_{n+m} = j \mid X_n = k, X_0 = i) P(X_n = k \mid X_0 = i)$$

$$= \sum_{k \in S} P(X_{n+m} = j \mid X_n = k) P(X_n = k \mid X_0 = i)$$

(Markov property)

$$= \sum_{k \in S} P(X_m = j \mid X_0 = k) P(X_n = k \mid X_0 = i)$$

(time-invariant property)

$$= \sum_{k \in S} P_{kj}^{(m)} P_{ik}^{(n)}$$

$\Rightarrow$ Forward equ: $P_{ij}^{(n+1)} = \sum_k P_{ik}^{(n)} P_{kj}$ for $n = 1, 2, \cdots$

and.

Backward equ: $P_{ij}^{(n+1)} = \sum_k P_{ik} P_{kj}^{(n)}$ for $n = 1, 2, \cdots$

→ single out final step and initial state i fixed.

This is useful when we want to know $P_{ij}^{(n)}$ for particular i but all values of j.

→ single out the change from the initial state i and has the final state j fixed.

This is useful when we want to know $P_{ij}^{(n)}$ for a particular j but all values of i.

In form of Matrix $P^{(n)} = \{ P_{ij}^{(n)} \} \in R^{S \times S}$

Forward equation: $P^{(n+1)} = P^{(n)} \cdot P$

where $P = P^{(1)}$: one-step transition Matrix.

Backward equation: $P^{(n+1)} = P \cdot P^{(n)}$

so $P^{(n)} = P^n$: Matrix Power.

Assume X_0 has the distribution $\vec{\mu}_0$ row vector.

$$\vec{\mu}_0 = [\mu_{01}, \mu_{02}, \cdots \mu_{0S}]$$

Let $\vec{\mu}_n$ be the marginal distribution of X_n

$$\mu_{nj} = P(X_n = j) = \sum_i P(X_n = j, X_0 = i)$$
$$= \sum_i P(X_n = j \mid X_0 = i) P(X_0 = i) = \sum_i \mu_{0i} P_{ij}^{(n)}$$

In Matrix form: $\vec{\mu}_n = \vec{\mu}_0 \mathbb{P}^n$

Q1: Any property of P ?

Q2: How to calculate P^n ?

Q3: Does $\lim_{n\to\infty} \vec{\mu}_n$ exist? If exists, what is it?

Q1: All rows of P sum to 1. $\lambda(P) = 1$

so $P \mathbb{1} = \mathbb{1}$ $\Rightarrow$ $\boxed{P \text{ has the right eigenvector of } \mathbb{1} \text{ with eigenvalue } 1.}$

$$\mathbb{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

All other eigenvalues of P $|\lambda(P)| \leq 1$ (why?) (power Method). // $\lambda(P)$ are all within unit circle.

Q2: Assume P has eigen-decomposition.

$P = U \Lambda U^{-1}$ Note $\lambda_1 = 1 \geq |\lambda_2| \geq \cdots |\lambda_s|$.

and $U = [\,\vdots \mid \cdots\,]$

$$P^n = \underbrace{U \Lambda U^{-1}}_{=P} \underbrace{U \Lambda U^{-1}}_{=P} \cdots \underbrace{U \Lambda U^{-1}}_{=P}$$

$$= U \Lambda^n U^{-1}$$

Q3: This Q is tricky. Let's assume the Limit exists. We will come back to answer when the limit exists later.

Denote $\lim\limits_{n \to +\infty} \vec{\mu}_0 P^n = \vec{\pi}$

a bit non-rigorious here. we reach the stage that $\vec{\pi} = \vec{\pi} P$. we call $\vec{\pi}$ as stationary distribution/ invariant distribution/ equilibrium distribution.

$\boxed{\text{once we enter stationary, we will never leave}}$

Def: A probability vector $\vec{\pi}$ on MC is called stationary distribution of a stochastic matrix P if $\vec{\pi} P = \vec{\pi}$, i.e. $\pi_i = \sum_j \pi_j P_{ji}$ for all i and $\sum_i \pi_i = 1$, $\pi_i \geq 0$.

also called global balance equations.

$$\boxed{P_{ii} = 1 - \sum_{j \neq i} P_{ij}}$$

$$\pi_i = \sum_{j \neq i} \pi_j P_{ji} + \pi_i P_{ii}$$

$$\Longleftrightarrow \pi_i (1 - P_{ii}) = \sum_{j \neq i} \pi_j P_{ji}$$

$$\Longleftrightarrow \underbrace{\pi_i \sum_{j \neq i} P_{ij}}_{\text{prob of being in state } i \times \text{ the net flow out of state } i} = \underbrace{\sum_{j \neq i} \pi_j P_{ji}}_{\text{prob of being in each other state } j \times \text{ the net flow from that state into } i.}$$

Ex 1: $P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$

$$[\pi_0, \pi_1] \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} = [\pi_0, \pi_1]$$

$$\Rightarrow \begin{array}{l} (1-\alpha)\pi_0 + \beta\pi_1 = \pi_0 \\ \alpha\pi_0 + (1-\beta)\pi_1 = \pi_1 \end{array} \Rightarrow \pi_0 = \frac{\beta}{\alpha}\pi_1.$$

With $\pi_0 + \pi_1 = 1$ we have

$$\frac{\beta}{\alpha}\pi_1 + \pi_1 = 1 \Rightarrow \pi_1 = \frac{\alpha}{\alpha+\beta}.$$

then $[\pi_0, \pi_1] = \left[\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right]$ $\Leftarrow$ stationary distribution.

Ex 2:

Gambler's Ruin. $P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ with $p+q=1$

$$\pi_\alpha = [\alpha, 0, 0, 0, 1-\alpha]$$

$$\pi_\alpha P = \pi_\alpha \text{ for any } \alpha \in [0,1]$$

so it has an uncountable number of stationary distribution. // why?

Q3.1: When will a MC has a stationary distribution

Q3.2: how to find a stationary distribution.

Q3.3: When the stationary distribution will be unique.

We will answer this question in later lectures.

---

proof of $|\lambda(P)| \le 1$.

for any probability vector $\vec{\pi}_0$

$\vec{\pi}_n = \vec{\pi}_0 P^n$ must satisfies $\pi_{ni} \geq 0$ and $\|\pi_n\|_1 = 1$.

If one of λ, $|\lambda_j(P)| > 0$

$P^n = U \Lambda^n U^{-1}$ $\quad \vec{\pi}_0 = \sum_{i=1}^{N} a_i \vec{u}_i$

then $\vec{\pi}_n = \vec{\pi}_0 P^n = \sum_{i=1}^{N} a_i P^n \vec{u}_i = \sum_{i=1}^{N} a_i \lambda_i^n \vec{u}_i$

$\|\vec{\pi}_n\|_1 \geq c|\lambda_j|^n$ as $n \to +\infty$, $\|\vec{\pi}_n\|_1$ will eventually be larger than 1 which contradicts $\|\vec{\pi}_n\|_1 = 1$.

Another important question:

How to simulate Markov chain.

Given: transition matrix P and initial state $X_0 = i_0$.

Algorithm: simulate MC to first T steps.

① set initial value $X_0 = i_0$ and $n = 0$.

② sample a random variable Y follows the multinormial distribution. P_{i_n}

i.e. $Y \sim P_{i_n}$

i_{n-1} th row of P matrix.

and set $X_{n+1} = \underline{y}$, $n \leftarrow n+1$

③ If $n < T$, set $i_n = X_n$, and repeat step 2..

Otherwise stop.

Ex: In Gambler's Ruin.

$$P = \begin{array}{c} \\ \end{array}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{array}{l} \leftarrow \$0 \\ \leftarrow \$1 \\ \leftarrow \$2 \\ \leftarrow \$3 \\ \leftarrow \$4. \end{array}$$

$n=0$: $X_0 = \$1.$

$Y \sim [q, 0, p, 0, 0]$

$y = \$2.$

$n=1$: $X_1 = \$2.$

$Y \sim [0, q, 0, p, 0]$

$y = \$1.$

$n=2$: $X_2 = \$1$ and so on.