Announcements:
- HW1 is online, due next Wed in class
- OH: Mon 3:50-4:50 in BUCH A102, Fri 11-12 in my office, or by appointment

Recall $X_0, X_1, X_2, \dots$ is a Markov chain if

$$P(X_{n+1} = x_{n+1} \mid X_n = x_n, \dots, X_0 = x_0) = P(X_{n+1} = x_{n+1} \mid X_n = x_n)$$

$$\forall x_0, x_1, \dots, x_{n+1} \in S \rightarrow \text{state space.}$$

Examples: 4.1 – 4.7 in Ross book.

Definition:
- A Markov chain $(X_n)_{n\geq 0}$ is homogeneous if $\forall x, y \in S$, $P(X_{n+1} = x \mid X_n = y)$ is the same for all n.
- Given a homogeneous MC, by indexing the states
  $S = \{s_1, s_2, \dots, s_i, \dots\}$
  we can define the transition matrix $\underset{\sim}{P}$ of $(X_n)_{n\geq 0}$ as
  $(\underset{\sim}{P})_{ij} = p_{ij} = P(X_{n+1} = s_j \mid X_n = s_i)$

Note: $\underset{\sim}{P}$ could have infinitely many rows and columns

Remark: When S is infinite, we generalize the classical concept of a finite matrix to a matrix with infinitely many rows and columns

Example (4.2 in Ross): $X_n \in \{0, 1\}$

Communication system with 2 states 0 and 1

Probability that message at next step is unchanged is p $(0 < p < 1)$, probability changed is $1-p$.

Then, $P(X_{n+1} = 0 \mid X_n = 0) = p = P(X_{n+1} = 1 \mid X_n = 1)$

$P(X_{n+1} = 0 \mid X_n = 1) = P(X_{n+1} = 1 \mid X_n = 0) = 1-p.$

$$\Rightarrow \quad \underset{\sim}{P} = \begin{array}{cc|c} 0 & 1 & \\ \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix} & & \begin{array}{c} 0 \\ 1 \end{array} \end{array}$$

Properties of the transition matrix $\underset{\sim}{P}$:

(i) $\forall i,j, \quad 0 \le p_{ij} \le 1$ $\rightsquigarrow$ p_{ij} is a probability.

(ii) $\forall i, \quad \sum_j p_{ij} = 1$

$$\sum_j P(X_{n+1} = s_j \mid X_n = s_i) = 1$$

given we are in state s_i, we have to go to one of the states in S w.p. 1.

A matrix satisfying (i) and (ii) is a STOCHASTIC MATRIX.

Exercise: Suppose the distribution of the first symbol X_0 is $\mu = (\mu_0, \mu_1)$. What is the distribution of X_1? of X_2? of X_n?

Answer: $P(X_1 = 0) = P(X_1 = 0, X_0 = 0) + P(X_1 = 0, X_1 = 1)$

$= P(X_1 = 0 \mid X_0 = 0)\, P(X_0 = 0) + P(X_1 = 0 \mid X_1 = 1)\, P(X_1 = 1)$

$= p_{00}\, \mu_0 + p_{10}\, \mu_1$

$P(X_1 = 1) = P(X_1 = 1 \mid X_0 = 0)\, P(X_0 = 0) + P(X_1 = 1 \mid X_0 = 1)\, P(X_1 = 1)$

$= p_{01}\, \mu_0 + p_{11}\, \mu_1$

$$\big(P(X_1=0),\ P(X_1=1)\big) = \big(p_{00}\mu_0 + p_{10}\mu_1,\ p_{01}\mu_0 + p_{11}\mu_1\big)$$

$$= (\mu_0, \mu_1)\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

$$= \mu \underset{\sim}{P}$$

Similarly, the distr. of X_2 is

$$\big(P(X_2=0), P(X_2=1)\big) = \big(P(X_1=0), P(X_1=1)\big)\underset{\sim}{P}$$

$$= (\mu \underset{\sim}{P})\underset{\sim}{P} = \mu \underset{\sim}{P}^2.$$

The distribution of X_n is $\mu \underset{\sim}{P}^n$.

Proposition: If at time n, the distribution of X_n is $\mu = (\mu_1, \dots, \mu_N)$, then the distr of X_{n+1} is $\mu \cdot \underset{\sim}{P}$ because:

$$P(X_{n+1}=s_j) = \sum_{k\in S} \underbrace{P(X_{n+1}=s_j \mid X_n = s_k)}_{p_{kj}} \underbrace{P(X_n = s_k)}_{\mu_k}$$

$$= \sum_{k\in S} \mu_k p_{kj} = (\mu \cdot \underset{\sim}{P})_j$$

$$\Rightarrow \big(P(X_{n+1}=s_1), \dots, P(X_{n+1}=s_N)\big) = \mu \cdot \underset{\sim}{P}.$$

Also, distr. of X_{n+2} is $\mu \cdot \underset{\sim}{P}^2$, etc.

Thus, if X_0 has distr μ, then X_n has distr $\mu \cdot \underset{\sim}{P}^n$.

Definition: Let $(X_n)_{n\geq 0}$ be a Markov chain with transition matrix P, and let μ be a distribution on S. We say that μ is a stationary distribution of $(X_n)_{n\geq 0}$ if

$$\mu \underset{\sim}{P} = \mu.$$

Proposition: If the distribution μ_n of X_n converges to a distr. μ, then μ is stationary.

Proof: $\mu_{n+1} = \mu_n \cdot \underset{\sim}{P} \Rightarrow \mu = \mu \cdot \underset{\sim}{P} \Rightarrow \mu$ is stationary.

$$\begin{array}{ccc} \mu_{n+1} & = & \mu_n \cdot \underset{\sim}{P} \\ \downarrow{\scriptstyle n\to\infty} & & \downarrow{\scriptstyle n\to\infty} \\ \mu & & \mu \end{array}$$

Example: If $\underset{\sim}{P} = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix}$, then what is a stationary distr?

$(\mu_0\ \mu_1)\underset{\sim}{P} = (p\mu_0 + (1-p)\mu_1, (1-p)\mu_0 + p\mu_1)$ has to equal to (μ_0, μ_1)

Solving for μ_0 and μ_1 ... we get $(\mu_0, \mu_1) = (\frac{1}{2}, \frac{1}{2})$.

议门: