Note 10 - Feb 07

Review:

Definition: DTMC

1. discrete state space
2. Markov property: present state

$$P(X_{n+1}=x_{n+1}|X_n=x_n,...,X_0=x_0)=P(X_{n+1}=x_{x_1}|X_n=x_n)$$

If $P(X_{n+1}=j|X_n=i)=P(X_1=j|X_0=i)$: time homogeneous (default setting)

$$\begin{aligned} P_{ij} &:= P(X_1=j|X_0=i) \\ &= P(X_{n+1}=j|X_n=i) \end{aligned}$$

$p=\{P_ij\}_{i,j \sin S}$: (one-step) transition (probability) matrix

Property:

- $P_{ij}\geq 0$, $i,j\in S$
- $\sum_{j\in S}P_{ij}=1$, $\forall i\in S$. (row sum)

4. Stochastic Processes (cont'd)

4.2 Chapman-Kolmogorov equations

Q: Given the (one-step) transition matrix, $P=\{P_{ij}\}_{i,j\in S}$, how can we decide the n-step transition probability

$$\begin{aligned} P_{ij}^{(n)} &:= P(X_n=j|X_0=i) \\ &= P(X_{n+m}=j|X_m=i), \quad m=0,1,... \end{aligned}$$

As a special case, let us start with $P_{ij}^{(2)}$ and their collection $p^{(2)}=\{P_{ij}^{(2)}\}_{i,j\in S}$ (also a square matrix, same dimension as P)

Condition on what happens at time 1:

$$\begin{aligned} P_{ij}^{(2)} &= P(X_2=j|X_0=i) \\ &= \sum_{j\in S} P(X_2=j|X_0=i,X_1=k)\cdot P(X_1=k|X_0=i) \quad \text{conditional law of total probability} \end{aligned}$$

4.2.1 Conditional Law of total probability

$$
\begin{aligned}
&P(X_2 = j|X_0 = i) \\
&= \sum_{k\in S} P(X_2 = j, X_1 = k|X_0 = i) \\
&= \sum_{k\in S} \frac{P(X_2 = j, X_1 = k, X_0 = i)}{P(X_0 = i)} \\
&= \sum_{k\in S} \frac{P(X_2 = j, X_1 = k, X_0 = i)}{P(X_1 = k, X_0 = i)} \cdot \frac{P(X_1 = k, X_0 = i)}{P(X_0 = i)} \\
&= \sum_{k\in S} P(X_2 = j|X_0 = i, X_1 = k) \cdot P(X_1 = k|X_0 = i)
\end{aligned}
$$

continue on $P_{ij}^{(2)}$

$$
\begin{aligned}
P_{ij}^{(2)} &= P(X_2 = j|X_0 = i) \\
&= \sum_{j\in S} P(X_2 = j|X_0 = i, X_1 = k) \cdot P(X_1 = k|X_0 = i) \quad \text{conditional law of total probability} \\
&= \sum_{k\in S} P(X_2 = j|X_1 = k) \cdot P(X_1 = k|X_0 = i) \\
&= \sum_{k\in S} P(X_1 = j|X_0 = k) \cdot P(X_1 = k|X_0 = i) \\
&= \sum_{k\in S} P_{ik} \cdot P_{kj} \\
&= (P \cdot P)_{ij}
\end{aligned}
$$

Thus, $p^{(2)} = P \cdot P = p^2$

Using the smae idea, for $n, m = 0, 1, 2, 3...$:

$$
\begin{aligned}
P_{ij}^{(n+m)} &= P(X_{n+m} = j|X_0 = i) \\
&= \sum_{k\in S} P(X_{n+m} = j|X_0 = i, X_m = k) \cdot P(X_m = k|X_0 = i) \\
&= \sum_{k\in S} P(X_{n+m} = j|X_m = k) \cdot P(X_m = k|X_0 = i) \quad \text{Markov property} \\
&= \sum_{k\in S} P(X_n = j|X_0 = k) \cdot P(X_m = k|X_0 = i) \\
&= \sum_{k\in S} p_{ik}^{(m)} \cdot P_{kj}^{(n)} \\
&= (p^{(m)} \cdot p^{(n)})_{ij} \\
&\Rightarrow p^{(n+m)} = p^{(m)} \cdot p^{(n)} \qquad (*)
\end{aligned}
$$

By definition, $p^{(1)} = p$

- $\Rightarrow p^{(2)} = p^{(1)} \cdot p^{(1)} = p^2$
- $\Rightarrow p^{(3)} = p^{(2)} \cdot p^{(1)} = p^3$
- $\cdots\cdots\cdots$
- $\Rightarrow p^{(n)} = p^n$

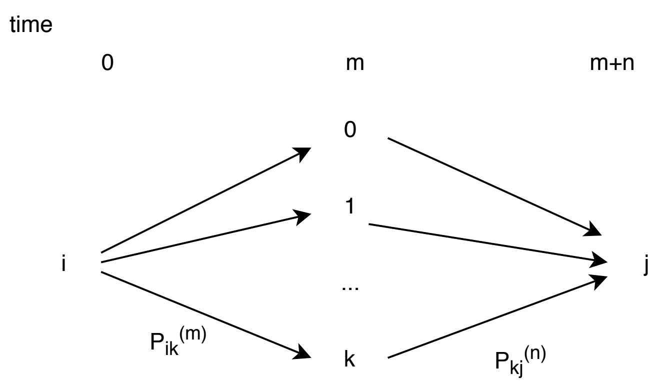
Note:

- n from $p^{(n)}$: n-step transition probability matrix
    - $p^{(n)} = \{p_{ij}^{(n)}\}_{i,j \in S}$
      $p_{ij}^{(n)} = P(X_n = j | X_0 = i)$
- n from p^n: n-th power of the (one-step) transition matrix
    - $p^n = p \cdot ... \cdot p$
      $p = \{P_{ij}\}_{i,j \in S}$
      $p_{ij} = P(X_1 = j | X_0 = i)$

$(*)$ is called the **Chapman-Kolmogorov equations** (c-k equation). Entry-wise:

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

Intuition:



"Condition at time m (on X_m) and sum p all the possibilities"

4.2.2 Distribution of X_n

So far, we have seen transition probability $P_{ij}^{(m)} = P(X_n = j | X_0 = i)$. This is not the probability $P(X_n = j)$. In order to get this distribution, we need the information about which state the Markov chain starts with.

Let $\alpha_{0,i} = P(X_0 = i)$. The row vector $\alpha_0 = (\alpha_0, 0, \alpha_0, 1, ...)$ is called the **initial distribution** of the Markov chain. This is the distribution of the initial state X_0

Similarly, we define distribution of X_n: $\alpha_n = (\alpha_n, 0, \alpha_n, 1, ...)$ where $\alpha_{n,i} = P(X_n = i)$

Fact: $\alpha_n = \alpha_0 \cdot p^n$

Proof:

$$\forall j \in S$$

$$\begin{aligned}
\alpha_{n,j} &= P(X_n = j) \\
&= \sum_{i \in S} P(X_n = j | X_0 = i) \cdot P(X_0 = i) \\
&= \sum_{i \in S} \alpha_{0,i} \cdot P_{ij}^{(n)} \\
&= (\alpha_0 \cdot P^{(n)})_j = (\alpha_0 \cdot p^n)_j
\end{aligned}$$

$$\Rightarrow \alpha_n = \alpha 0 \cdot p^n$$

- α_n: distribution of X_n
- α_0: initial distribution
- p^n: transition matrix

Remark: The distribution of a DTMC is completely determined by two things:

- the initial distribution α_0 (row vector), and
- the transition matrix p (square matrix)

4.3 Stationary distribution (invariant distribution)

Definition: A probability distribution $\pi = (\pi_0, \pi_1, ...)$ os ca;;ed a **stationary distribution**(invariant distribution) of the DTMC $\{X_n\}_{n=0,1,...}$ with transition matrix P, if :

1. $\underline{\pi} = \pi \cdot P$
2. $\sum_{i \in S} \pi_i = 1 (\Leftrightarrow \underset{\leftarrow}{\pi} \cdot \mathbb{1})$. ($\mathbb{1}$: a column of all 1's)

Why such $\underline{\pi}$ is called stationary/invariant distribution?

$$\sum_{i \in S} \pi_i = 1, \pi_i \geq 0, i = 0, 1, ... \Rightarrow \text{distribution}$$
$$\underline{\pi} = \pi \cdot P \Rightarrow \text{invariant/stationary.}$$

Assume the MC starts from the initial distribution $\alpha_0 = \underline{\pi}$. hen the distribution of X_1 is

$$\alpha_1 = \alpha_0 \cot P = \underline{\pi} \cdot P = \underline{\pi} = \alpha_0$$

The distribution of X_2:

$$\alpha_2 = \alpha_0 \cdot P^2 = \underline{\pi} \cdot P \cdot P = \underline{\pi} \cdot P = \underline{(}\pi) = \alpha_0$$

$$\cdots\cdots$$

$$\alpha_n = \alpha_0$$

Thus, if the MC starts from a stationary distribution, then its distribution will not change over time.