Lecture 8

Math 50051, Topics in Probability Theory and Stochastic Processes

Markov Processes

Markov processes play an important role in derivative asset pricing. Our discussion will be mostly in discrete time, and we will try to motivate some important aspects of stochastic processes and will clarify some notions that will be used in dealing with continuous time models for interest rate derivatives. It is quite important that the process one is modeling in finance is a Markov process. The Feynman-Kac theorem that we will see later (maybe next semester) will be valid only for such processes. However, it can be shown that short-term interest rate processes are not, in general, Markov. This imposes limitations on the numerical methods that can be applied for short rate processes. Let's consider a discrete-time, discrete space sp $\{X(t), t \in N\}$, with values (the state space) $V = \{0, 1, 2, \ldots, N\}$. To characterize the distribution of the stochastic process X it is sufficient to have **all** the finite dimensional distributions, ie we want the values of

$$P(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n)$$

for every n and every finite sequence of states $(i_0, \cdots, i_n)$.

Remark: 1) $i_0, i_1, \cdots, i_n \in V$
2) When we say $P(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n)$ we mean $P(\{X_0 = i_0\} \cap \{X_1 = i_1\} \cap \cdots \cap \{X_n = i_n\})$
3) Knowing **all** $P(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n)$ is equivalent to knowing the **initial distribution**, denoted in general by π, ie

$$\pi_1^{(0)} = P(X_0 = i), \ i \in V$$

and the **transition probabilities**

$$P(X_n = i_n | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}), \quad \text{for all } i_0, i_1, \ldots, i_n \in V, \ \ n = 1, 2, \ldots,$$

Why are these two things equivalent?

<u>Discrete time Markov Chains</u> Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(X_n)_n$ be a discrete time stochastic process with state S (discrete) (X_n takes values in S). We say that X is a <u>Markov Chain</u> on S if for all $n \in \mathbb{N}$ and $s \in S$

$$P(X_{n+1} = s | X_0, ..., X_n) = P(X_{n+1} = s | X_n) \qquad *$$

$*$ is called the <u>Markov property</u> of X_n, $n \in \mathbb{N}$. Because S is discrete, * is equivalent to:

$$P(X_{n+1} = s_{n+1} | X_0 = s_0, X_1 = s_1, ...X_n = s_n) = P(X_{n+1} = s_{n+1} | X_n = s_n) \tag{5}$$

$$for\ all\ n \in \mathbb{N}\ and\ for\ all\ s_0, s_1, ...s_{n+1} \in S$$

<u>Remark:</u> The assumption of Markovness has more than just theoretical relevance in asset pricing. The Markov property states that to make predictions of the behavior of a system in the future, it is sufficient to consider only the present state of the system and not the past history. That is, the state of the system is important but not how it arrived at that state.

The relevance

1) How does these notions help a market practitioner? Suppose that X_t represents a variable such as instantaneous spot rate r_t. Then, assuming that r_t is Markov means that the expected future behavior of r_{t+s} depends only on the latest observation and using this property one can captures the dynamics of r_t, the interest rate. But if interest rates are not Markovian, since the conditional means and variances of the spot rate could potentially depend on observations other than the immediate past one could not determine the dynamics of r, so at least for interest rates derivatives the Markovian property is important.

2) We will see that, although two processes could be jointly Markov, when we model one of these processes in a univariate setting, it will, in general, cease to be a Markov process. The relevance of this can best be described in fixed income. There, a central concept is the yield curve. The classical approach attempts to model yield curve using a single interest rate process, such as r_t discussed above. The more recent Heath-Jarrow Merton approach, consistent with Black-Scholes philosophy models it using k separate forward rates, which are assumed to be Markov, jointly. But, as we will see, the univariate dynamics of one element of $k-$dimensional Markov process, will, in general, not be Markov.

Now, let's go back to our definitions.

Remark: We see that a $DTMC$ is completely characterized iff we have

(a) The initial distribution $\pi^{(0)}$

(b) The one-step transition probabilities:

$$P[X_{n+1} = j | X_n = i] \ \ \forall i, j \in S$$

for **all** n.

The MC is called time-homogeneous if

$$P(X_{n+1} = j | X_n = i) = P(X_n = j | X_{n-1} = i) = ... = P(X_1 = j | X_0 = i) = p_{ij} \ \ for\ all\ n.$$

Hence, for time homogeneous MC (a) and (b) are transformed into

(a') The initial distribution $\pi^{(0)}$

(b') p_{ij} for all $i, j \in S$.

p_{ij} are called transition probabilities from state i to state j and the matrix $P = (p_{ij})_{i,j \in S}$ is called the one-step transition matrix of the chain X_n.

Observed that:

(i) $0 \leq p_{ij} \leq 1$ (because p_{ij} are just some probabilities)

(ii) $\sum_{j \in S} p_{ij} = 1$

Indeed

$$\sum_{j \in S} p_{ij} = \sum_{j \in S} P(X_{n+1} = j | X_n = i)$$

$$= \sum_{j \in S} \frac{P(X_{n+1} = j, X_n = i)}{P(X_n = i)} = \frac{\sum_{j \in S} P(X_{n+1} = j, X_n = i)}{P(X_n = i)}$$

because the events $\{X_n = j\}$ are disjoint.

$$= \frac{P(\cup_{j \in S}\{X_{n+1} = j, X_n = i\})}{P(X_n = i)} = \frac{P(\cup_{j \in S}\{X_{n+1} = j\} \cap \{X_n = i\})}{P(X_n = i)}$$

$$= \frac{P(\Omega \cap \{X_n = i\})}{P(X_n = i)} = \frac{P(X_n = i)}{P(X_n = i)} = 1$$

This is because the process must be in one and only one state at time $n + 1$.

A <u>matrix</u> with properties (i) and (ii) is called **stochastic**.

The sum $\sum_{i \in S} p_{i,j}$ could take any nonnegative value. If we also have $\sum_{i \in S} p_{i,j} = 1$ for all j the matrix P is called **doubly stochastic**

<u>Examples:</u>

1) On any given day you are either H(happy), S(sad), or N(neutral). If you are H today, you will be H, S, N tomorrow with probabilities 0.5, 0.4, 0.1. If you are sad today, you will be H, S, N tomorrow with probabilities 0.3, 0.4, 0.3. If you are neutral today, then you will be H, S, N tomorrow with probabilities 0.2, 0.3, 0.5. let $X_n = \textit{your mood on the n'th day}$. What can you say about $\{X_n, n \geq 0\}$?

2) (Turning a non-Markov Chain to a MC) Suppose that whether or not it rains today depends on previous weather conditions through the last two days:

— If it has rained the past two days, then it will rain tomorrow with probability 0.7.

— If it rained today but not yesterday, then it will rain tomorrow with probability 0.5.

— If it rained yesterday but not today, tomorrow it will rain with probability 0.4.

— If it has not rained in the past two days, then it will rain tomorrow with probability 0.2.

Now let $\{X_n, n \geq 0\}$ be such that

$$X_n = \begin{cases} 1 & if\ it\ rains\ on\ day\ n \\ 0 & otherwise \end{cases}$$

What can you say about X_n? Now, let Y_n be defined as:

$$Y_n = \begin{cases} 1 \text{ if it rained on days } n, n-1 \\ 2 \text{ if it rained on days } n \text{ but not on } n-1 \\ 3 \text{ if it rained on days } n-1 \text{ but not on } n \\ 4 \text{ if it didn't rain on both } n \text{ and } n-1 \end{cases}$$

What can you say about Y_n?

More examples of MC

1) Let $S = \mathbb{Z}$ and (X_n) be a sequence of iid random variables with $P(X_1 = 1) = p$ and $P(X_1 = -1) = 1 - p$ (X_n tells you if you take a step to the left or to the right at time $t = n$). Define $Y_n = \sum_{i=1}^{n} X_i$ (Y_n tells you where you are at time $t = n$) and let $Y_0 = 0$. Show that Y_n is a Markov Chain and find its transitional probabilities. Y_n is called a <u>random walk on $\mathbb{Z}$ starting at 0</u>. Replacing $Y_0 = 0$ with $Y_0 = i$, we get a random walk that starts at i.

2) If gambler wins \$1 with probability p and loses \$1 with probability $1 - p$ at each play of the game. He quits when he goes broke or he attains a fortune of $\$N$. Let X_n be the gambler's fortune at time n. Then what can you say about $\{X_n, n \geq 0\}$? $\{X_n, n \geq 0\}$ is also called a random walk but on $S = \{0, 1, ..., N\}$ absorbed at 0 and N (absorbing barriers).

Transition probabilities in n steps Consider the case when the process moves from state i to state j in n steps. Of course, a first question to ask is what is the probability that a MC that is in state i will be in state j after n additional transitions. These transition probabilities in n steps, denoted by $p_{i,j}^{(n)} = P(X_{m+n} = j | X_m = i)$, for $m, n, i, j \geq 0$ are recorded in a matrix denoted by $P^{(n)}$, named **n-step transition matrix**. Observe that these probabilities do not depend on m, ie they are the same for any m. Indeed,

$$p_{i,j}^{(n)} = P(X_{m+n} = j | X_m = i) = P(X_n = j | X_0 = i),$$

because of the time homogeneity.

Turns out these matrices have some nice properties:

Theorem: Chapman-Kolmogorov equation Suppose that X_n is an $S-$valued Markov chain with $n-$step transition probabilities $p_{i,j}^{(n)}$. Then for all $k, n \in N$ we have

$$p_{i,j}^{(m+n)} = \sum_{S} p_{i,s}^{(m)} p_{s,j}^{(n)}, \quad \text{for all } i, j \in S$$

where $p_{i,j}^{(0)} = 1$ if $i = j$ and is 0 otherwise.

Remark: This implies that $P^{(m+n)} = P^{(m)} P^{(n)}$, and even more

$$P^{(n)} = P^n$$

Remark: If we are interested in unconditional probabilities associated to MC then we need to know the initial distribution, $\pi^{(0)}$, because

$$P(X_n = j) = \sum_{s \in S} p_{s,j}^{(n)} \pi_s^{(0)}.$$

Example: Suppose that if it rains today then it'll rain tomorrow with probability .7 and if it doesn't then it'll rain tomorrow with probability .4. What's the probability it will rain two days from now given that is raining today? If we know that the probability that will rain today is .4, what is the probability that it will rain in two days from now?

Classes and states We say state j is **accessible** from state i or that state i **communicates** with state j if $p_{i,j}^{(n)} > 0$ for some $n \geq n$, ie the chain will visit state j if it starts in state i in n steps, for some n positive. Two states i and j **intercommunicate** if they are both accessible or if both communicate to each other.

Notation:
$i \to j$ j is accessible, or i communicates with j.
$i \leftrightarrow j$ i and j intercommunicate.

Remark: If two states do not intercommunicate then either

$$p_{i,j}^{(n)} = 0 \quad \text{for all } n \geq 0$$

or

$$p_{j,i}^{(n)} = 0 \quad \text{for all } n \geq 0$$

Properties: Intercommunication is an equivalence relationship, meaning it verifies the following 3 properties

1) It is reflexive, ie $i \leftrightarrow i$.
2) It is symmetric, ie $i \leftrightarrow j$ implies $j \leftrightarrow i$;
3) It is transitive, ie, if $i \leftrightarrow j$ and $j \leftrightarrow k$ then $i \leftrightarrow k$.
Why is this true?

Two states that intercommunicate are said to be in the same **class**. If **all** states intercommunicate then we say that the MC is **irreducible**. If there are several classes in the MC we say the MC is reducible.

Examples:

Notation Denote by $f_{ii}^{(n)} = P[X_n = i, X_m \neq i, m = 1, 2, ..., n-1 | X_0 = i]$ – the probability that starting from i, the first return to i occurs at the n'th transition and by $f_{ii} = \sum_{n=0}^{\infty} f_{ii}^{(n)}$ – the probability that starting from i, the process will ever return to i. ($f_{ii}^{(0)} = 0$).

Definition State i is said to be recurrent iff $f_{ii} = 1$; State i is said to be transient iff $f_{ii} < 1$.

Consequences:

c1) State i is recurrent implies that starting in i, the process will revisit i infinitely many times with probability 1 (the M.P. and time homogeneity implies this.)

c2) i transient $\Longrightarrow P[X \text{ will ever return } i | X_0 = i] = f_{ii}$

$\Longrightarrow P[X\ will\ be\ in\ i\ for\ exactly\ n\ time\ epochs|X_0 = i] = f_{ii}^{n-1}(1-f_{ii})$. **Note** it is f_{ii}^{n-1} not $f_{ii}^{(n-1)}$.

$\Longrightarrow$ Given $X_0 = i$, the number of time epochs in state i is distributed $Geom(1 - f_{ii})$.

$\Longrightarrow E[number\ of\ times\ epochs\ in\ i|X_0 = i] = \frac{1}{1-f_{ii}} < \infty$.

$\Longrightarrow P[X\ revisits\ i\ \infty\ times\ |X_0 = i] = 0$.

Proposition: Let $X = \{X_n\}$ be a $M.C$ with state space S. Let $i \in S$, then
If $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$, then i is recurrent.
If $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$, then i is transient.

Why?

Corollary 1 In a finite state $M.C.$ not all states can be transient.

Proof: Assume $S = \{1, 2, ..., N\}$ and suppose all states are transient. Let $T_i =$ the time after which state i will never be revisited ($P[T_i < \infty] = 1, i \in S$). Then after time $T = max\{T_i\}$ no state will be revisited, with probability 1. But this is absurd.

Corollary 2: a) i recurrent & $i \longleftrightarrow j \Longrightarrow j$ is recurrent, and
i transient & $i \longleftrightarrow j \Longrightarrow j$ is transient.

b) All states in a finite irreducible $M.C.$ are recurrent.

Example 1

$$\begin{pmatrix} 0 & 0 & 1/3 & 2/3 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

This is a finite irreducible $M.C.$, so all states are recurrent.

Example 2

$$\begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 0 & 0 & 1/2 \end{pmatrix}$$

reducible and we have 3 classes $\{0, 1\}, \{2, 3\}$ (recurrent), $\{4\}$ (transient).