MATH 285 A 4/4/2007

CLASS 2

NOTETAKER : ALBERTO ROSSI
ASSISTANT NOTETAKER : AYELEN BANEGAS

ANNOUNCEMENTS :
- NO CLASS ON APRIL 25
- TEACHING ASSISTANT : VLADIMIR PESIC
  APM 6321
  OFFICE HOURS : MON 11.00 A.M. - 1.00 P.M.

QUICK REVIEW.

- LAW OF STOCHASTIC PROCESS
  (DISTRIBUTION OF A STOCHASTIC PROCESS)

- SINGLE RANDOM VARIABLE. (REAL VALUED)

PROBABILITY SPACE $(\Omega, \mathcal{F}, P)$ [IN THE BACKGROUND, AS USUAL]

$X : \Omega \longrightarrow \mathbb{R}$ (THE RANDOM VARIABLE IS A MAPPING FROM THE SAMPLE SPACE TO THE REAL LINE)

(FOR ANY INTERVAL OF THE FORM (a,b) : $-\infty < a < b < \infty$,

$$X^{-1}\left((a,b)\right) = \left\{ \omega \in \Omega : X(\omega) \in (a,b) \right\} \in \mathcal{F} \quad)$$

X AND P INDUCE A PROBABILITY MEASURE ON $\mathbb{R}$

$\mu\left((a,b)\right) = P\left(X \in (a,b)\right)$ for $-\infty < a < b < \infty$. WE CALL "μ" THE DISTRIBUTION (OR LAW) OF X.

- **Distribution Function**
  (of a Single Random Variable)

$$F_X(x) = P(X \leq x)$$
$$= P(X \in (-\infty, x])$$
$$= \mu((-\infty, x])$$

NOTE: F_X is RIGHT-CONTINUOUS

- **Stochastic Processes**

A collection of random variables indexed by time is what we call a stochastic process.

Formally:

$$X = \{X_t, t \in \mathbb{T}\}$$

For each t, X_t is a random variable taking values in a state space $(S, \mathcal{S})$

S: state space

$\mathcal{S}$: σ-algebra on the state space.

If we fix "t", we have one random variable X_t and we can look at its (individual) distribution.

• Finite Dimensional Distributions

For each set $\{t_1, \ldots, t_m\} \subset \mathbb{T}$, $(X_{t_1}, \ldots, X_{t_m})$ is an m-dimensional vector and specifying its distribution for each $\{t_1, \ldots, t_m\}$ gives the finite dimensional distributions of X.

The distribution of $(X_{t_1}, \ldots, X_{t_m})$ is determined by specifying

$$P\Big((X_{t_1}, \ldots, X_{t_m}) \in A_1 \times A_2 \times \ldots \times A_m \Big) \text{ for any } A_1, \ldots, A_m \in \mathcal{S}$$

and this gives a probability measure on S^m

Letting $(t_1, \ldots, t_m)$ run over all subsets of $\mathbb{T}$, we can determine a probability measure on $S^{\mathbb{T}}$ and the probability measure is the law or probability distribution of the process.

NOTE: In principle, in order to specify the distribution of the stochastic process, we would need to specify the finite dimensional distributions for each set $\{t_1, \ldots, t_m\} \subset \mathbb{T}$. Fortunately, we have clever ways of specifying a stochastic process other than writing down the finite dimensional distributions.

(The idea is to specify rules to ensure that the stochastic process is well defined)

End of introduction. We start now analyzing in detail one class of stochastic processes.

• MARKOV CHAINS (INFORMALLY)

THE TERM COMES FROM THE DISCRETE NATURE OF THE STATE SPACE

(FINITE OR COUNTABLY INFINITE)

(We could use NON-NEGATIVE INTEGERS TO LABEL STATES)

INFORMALLY: THE DISTRIBUTION OF WHAT WILL HAPPEN IN THE FUTURE DEPENDS ON THE PAST <u>ONLY THROUGH</u> THE CURRENT STATE.

NOTE: TIME HAS AN ORIENTATION

• MARKOV PROPERTY

For $\{X_t, t \in \mathbb{T}\}$,

$$P(\underbrace{X_t \in A}_{\text{FUTURE EVENT}} \mid \underbrace{X_m : m \leq s}_{\text{what happened in the past up to time } s}) = P(X_t \in A \mid \underbrace{X_s}_{\text{CURRENT STATE}})$$

$s < t$

NOTE THE IMPLICIT ORDERING OF THE TIME

WE WILL NOW PROCEED TO A MORE FORMAL TREATMENT OF DISCRETE TIME MARKOV CHAINS

Discrete Time Markov Chains

$\mathbb{T} = \mathbb{N} = \{0, 1, 2 \ldots\}$ (could also take it to be finite : i.e. $\mathbb{T} = \{0, 1, \ldots, N\}$)

S = discrete state space (either finite or countably infinite)

(can index by subset of non-negative integers)

$\mathcal{S}$ = all subsets of the state space S.

Definition:

With DTMC we refer to a stochastic process $\{X_m, m = 0, 1, 2 \ldots\}$ where X_m takes values in S for each m and

$$P(X_m = i_m \mid X_0 = i_0, X_1 = i_1, \ldots, X_{m-1} = i_{m-1}) = P(X_m = i_m \mid X_{m-1} = i_{m-1})$$

for all $i_0, i_1, \ldots, i_m \in S$, $m \geq 1$

Example:

Here is how we can rewrite the expression $P(X_2 = i_2 \mid X_0 = i_0)$ using Markov property.

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

$$= \frac{\sum_J P(X_2 = i_2, X_1 = J, X_0 = i_0)}{P(X_0 = i_0)}$$

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

uses Markov property

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

Throughout THE ANALYSIS, WE WILL ASSUME THAT MARKOV CHAINS ARE TIME HOMOGENEOUS.

i.e. $P(X_m = i_m \mid X_{m-1} = i_{m-1})$ DOES <u>NOT</u> DEPEND ON m.

- <u>TRANSITION "MATRIX"</u>

Let $P_{ij} = P(X_m = j \mid X_{m-1} = i) = P(X_1 = j \mid X_0 = i)$

(this "matrix" can HAVE INFINITELY MANY ROWS and COLUMNS

All ENTRIES ARE NON-NEGATIVE AND

$$\sum_j P_{ij} = \sum_{j \in S} P(X_1 = j \mid X_0 = i)$$

$$= P(X_1 \in S \mid X_0 = i) \quad \text{FOR ALL } i$$

$$= 1 \qquad (*)$$

To specify THE DISTRIBUTION OF A DISCRETE MARKOV CHAIN, we JUST NEED the TRANSITION PROBABILITY MATRIX P and THE INITIAL DISTRIBUTION λ:

$$\underbrace{\lambda(i)}_{\text{DISTRIBUTION ON THE STATE SPACE.}} = P(X_0 = i) \quad \text{for all } i \in S$$

NOTE: P^m = m^{th} POWER OF THE TRANSITION "MATRIX".

(*) (IT MAY SOUND WEIRD BECAUSE THE "MATRIX" AT HAND MAY HAVE INFINITE ROWS and COLUMNS, BUT IT IS WELL BEHAVED BECAUSE THE ELEMENTS OF EACH ROW SUM UP TO 1) e.g. $P^2_{ij} = \sum_{k \in S} P_{ik} P_{kj}$

$$(P^n)_{ij} = P(X_n = j \mid X_0 = i) \quad \text{CHECK THIS :}$$

HINT: DO IT FOR $n=2$ AND PROCEED BY INDUCTION.

- **FINITE DIMENSIONAL DISTRIBUTION** (DETERMINE LAW OF MARKOV CHAIN)

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \underbrace{P(X_n = i_n \mid X_0 = i_0, \dots, X_{n-1} = i_{n-1})} P(X_0 = i_0, \dots, X_{n-1} = i_{n-1})$$

$i_0, i_1, \dots, i_n$; $n = 0, 1, \dots$

$$\overset{\text{MARKOV PROPERTY}}{=} \underbrace{P(X_n = i_n \mid X_{n-1} = i_{n-1})}_{P_{i_{n-1} i_n}} P(X_0 = i_0, \dots, X_{n-1} = i_{n-1})$$

Proceeding recursively

$$\vdots$$

$$= (P_{i_{n-1} i_n} P_{i_{n-2} i_{n-1}} \cdots P_{i_0 i_1}) \lambda(i_0)$$

QUESTIONS OF INTEREST

① FINITE TIME HORIZON :: FIRST PASSAGE PROBLEMS

- HITTING TIME - DISTRIBUTION
- ABSORPTION PROBABILITY QUESTIONS
- ABSORPTION TIME QUESTIONS

② LONG-RUN BEHAVIOR : • FRACTION OF TIME SPENT IN A STATE (INFINITE-TIME HORIZON)

NOW EXAMPLE OF A PROCESS THAT IS <u>NOT</u> MARKOV.

TAKE $\{\xi_n\}_{n=1}^{\infty}$ i.i.d AND SUPPOSE THAT

$$P(\xi_n = +1) = P(\xi_n = -1) = P(\xi_n = 0) = \frac{1}{3}$$

$S_n = \sum_{i=1}^{n} \xi_i \qquad S_0 = 0$

NOTE: $\{S_n\}_{n=0}^{\infty}$ IS MARKOV

Now, Define $\quad X_n = \max\{S_0, S_1, \dots, S_n\}$

$P(X_4 = i_4 \mid X_0 = i_0, X_1 = i_1, X_2 = i_2, X_3 = i_3)$ IS THE PROBABILITY OF REACHING A NEW MAXIMUM GIVEN WHAT HAPPENED IN THE PAST. IT WILL NOT BE ENOUGH TO KNOW WHAT HAPPENED AT X_3, BECAUSE THE MAX DEPENDS ON ALL X_0, X_1, X_2, X_3.

LUCKILY, IN ORDER TO SHOW THAT A PROCESS IS NOT MARKOV, WE JUST NEED A COUNTER-EXAMPLE.

Need to show

$$P(X_4 = i_4 \mid X_0 = i_0, X_1 = i_1, X_2 = i_2, X_3 = i_3) \neq P(X_4 = i_4 \mid X_3 = i_3)$$

SAY FOR SOME i_0, i_1, i_2, i_3, i_4.