STOCHASTIC PROCESSES & TRANSITION PROBABILITIES

INTRODUCTION

In the previous lecture we learnt the basics of Chance Constrained Linear Programming (CCLP) and its application in reservoir operation. Another type of Explicit Stochastic Optimization (ESO) model is the Stochastic Dynamic Programming (SDP). Before dealing with SDP models, it is necessary to have an understanding on stochastic processes and transition probabilities which will be discussed in the present lecture.

STOCHASTIC PROCESSES

Random hydrological variables (e.g. rainfall, streamflow) are obtained as a sequence of observations of historic record, called a time series. In time series, the observations are ordered with reference to time. These observations are often dependent, i.e., r.v. at one time influences the r.v. at later times. Thus, time series is essentially observation of a single r.v.

A r.v. whose value changes with time according to some probabilistic laws is termed as a stochastic process. A time series is one realization of a stochastic process. A single observation at a particular time is one possible value a r.v. can take. A stochastic process is a time series of a r.v.

As stated above a single time series is one realization of a stochastic process. The statistical properties from one realization will be equal to those from another realization, only if the process is stationary. In other words, the probability distribution of a stationary process does not vary with time.

MARKOV PROCESSES AND MARKOV CHAINS

In water resources system modeling, most of the stochastic processes are often treated as Markov process. A process is said to be a **Markov process** (first order) if the dependence of future values of the process on the past values is completely determined by its dependence on the current value alone, i.e., a first order Markov process has the property

$$P\left[X_{t+1} / X_t, X_{t-1}, \ldots, X_0\right] = P\left[X_{t+1} / X_t\right]$$

Since the current value summarizes the state, it is often referred as the **state**. A Markov process whose state X_t takes only discrete values is termed **Markov chain**.

The assumption of a Markov Chain implies that the dependence of any hydrological variable in the next period on its current and all previous periods' values is completely described by its dependence on its current period value alone. The dependence of a r.v. in the next period on the current value is expressed in terms of transition probabilities.

TRANSITION PROBABILITIES

Consider a stream whose inflow Q is a stationary random variable. Transition probabilities measure the dependence of the inflow during period t+1 on the inflow during the period t. Transition probability P_{ij}^{t} is defined as the probability that the inflow during the period t+1 will be in the class interval j, given that the inflow during the period t lies in the class interval i.

$$P_{ij}^{t} = P[Q_{t+1} = j \,/\, Q_t = i]$$

where $Q_t=i$, indicates that the inflow during the period t belongs to the discrete class interval i. The historical inflow data is discretized first into suitable classes. Each inflow value in the historical data set is then assigned a class interval. P_{ij}^{t} are estimated from the number of times when the inflow in the period t belongs to class i and the inflow in period $t + 1$ goes to class j, divided by the number of times the inflow belongs to class i in period t.

Example

A sequence of inflows for 30 time periods are given in Table 1. Find the transition probabilities by discretizing the inflows into three intervals 0-2, 2-4 and 4-6.

Table 1

t	Q_t	t	Q_t	t	Q_t
1	2.4	11	1.6	21	4.8
2	2.3	12	1.3	22	4.1
3	1.5	13	2.4	23	5.5
4	1.1	14	1.6	24	5.9
5	2.1	15	3.4	25	3.2
6	2.4	16	2.6	26	4.3
7	4.2	17	3.5	27	5.3
8	4.6	18	2.6	28	3.2
9	5.1	19	1.4	29	1.2
10	3.2	20	4.5	30	4.6

Solution:

From the data,

Probability of being in first class 0-2, PQ_1= 7/30 = 0.23

Probability of being in second class 2-4, $PQ_2 = 12/30 = 0.40$

Probability of being in third class 4-6, $PQ_3 = 11/30 = 0.37$

The number of times a flow in interval j followed a flow in interval i is shown in the form of a matrix below. Here only 29 values are considered.

i \ j	1	2	3
1	2	3	2
2	4	6	2
3	0	3	7

Transition probabilities P_{ij} are then calculated by dividing each value by the sum of corresponding row values. Given an observed flow in an interval i in period t, the probabilities of being in one of the possible intervals j in the next period $t+1$ must sum to 1.

Transition probability Matrix.

i \ j	1	2	3
1	0.285	0.43	0.285
2	0.33	0.5	0.17
3	0.0	0.3	0.7

The sum of the probabilities in each row equals 1. Matrices of transition probabilities whose rows sum to 1 are also called **stochastic matrices or first-order Markov chains**.

STEADY STATE PROBABILITIES

One can compute the probability of observing a flow in any interval at any period in the future given the present flow interval, using the transition probability matrix. For example assume the flow in the current time period $t=1$ is in interval $i=2$. Following the above example, the probabilities, $PQ_{j,2}$, of being in any of the three intervals in the following time period $t=2$ are the probabilities shown in the second row of the matrix in the Table. The probabilities of being in an interval j in the following time period t=3 is the sum over all intervals i of the joint probabilities of being in interval i in period $t=2$ and making a transition to interval j in period $t=3$.

i.e., $$PQ_{j\,t+1} = \sum_i PQ_{i\,t}\; P_{i\,j}$$ for all intervals j and periods t.

This operation can be continued till N future time periods. After some time period, the flow interval probabilities as calculated above seem to be converging to the unconditional probabilities (calculated from the historical data). These are termed as the steady state probabilities.

As mentioned above, if the current period flow is in class i =2, then the probabilities are [0 1 0]. The steady state probabilities determined as per the above equation are given in the Table 2 which illustrates the results of such calculations for seven future periods.

Table 2. Steady state probabilities

Time period	Flow interval i		
	1	2	3
1	0	1	0
2	0.33	0.5	0.17
3	0.26	0.44	0.3
4	0.22	0.42	0.36
5	0.20	0.41	0.39
6	0.19	0.41	0.40
7	0.19	0.41	0.40

As the future time period t increases, the flow interval probabilities are converging to unconditional probabilities i.e., PQ_{it} will equal $PQ_{i,t+1}$ for each flow interval i. This indicates that as the number of time periods increases between the current period and that future time period, the predicted probability of observing a future flow in any particular interval at some time in the future becomes less and less dependent on the current flow interval.

Steady state probabilities can be determined by solving

$$PQ_j = \sum_i PQ_i \ P_{ij} \qquad \text{for all intervals } j.$$

and $$\sum_i PQ_i = 1$$