Stochastic Hydrology

Prof. P.P.Mujumdar

Department of Civil Engineering

Indian Institute of Science, Bangalore

Module No. # 07

Lecture No. # 34

Multivariate Stochastic Models – II

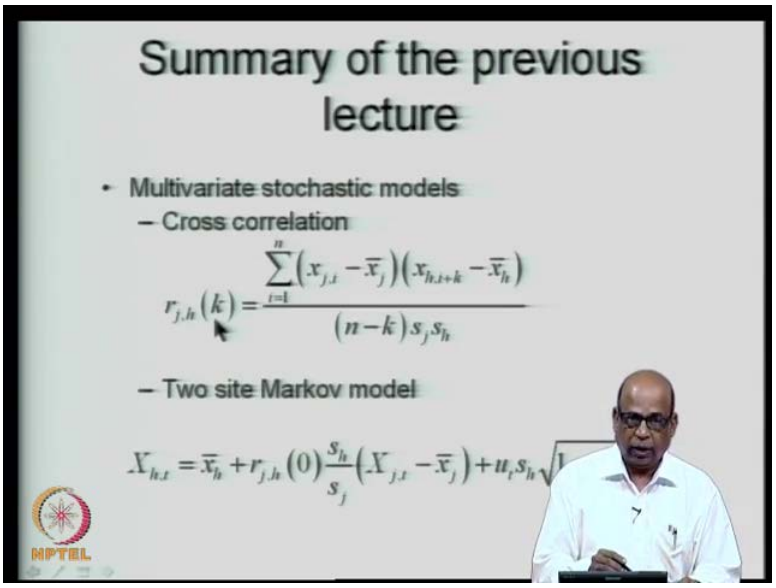
Good morning and welcome to this lecture number 34 of the course stochastic hydrology. We have been now discussing multivariate stochastic models, essentially to generate data on several sights simultaneously. Specifically, we deal with a stream flow generation. But, this can be also applied to other variables, other random variables, provided the data fits the particular assumptions that we make in this models.

So, in the last lecture, we looked at generation at two sites simultaneously. By simultaneously, I mean we also looked at the dependence of the flows at one site on the flows at other site. Distinguish this, between with what we had done earlier for single site models. In the single site data generation models, we considered the time series at a particular site and considered the serial correlations, which are also called as the autocorrelations.

In the two site model that we discussed in the previous lecture, what we essentially do is, we identify one as the key site depending on the length of the data, as well as on the quality of the data at that particular location. The other one, we treat is as a subordinate site to the key site that we have identified.

On the key site, we generate the data using the single site models. Then, on the subordinate site, we generate the data based on the covariance that exist between the data at the key site and the data at the subordinate site. So, this is what we essentially did in the previous lecture, where we talked about data generation simultaneously at two locations.

(Refer Slide Time: 02:28)



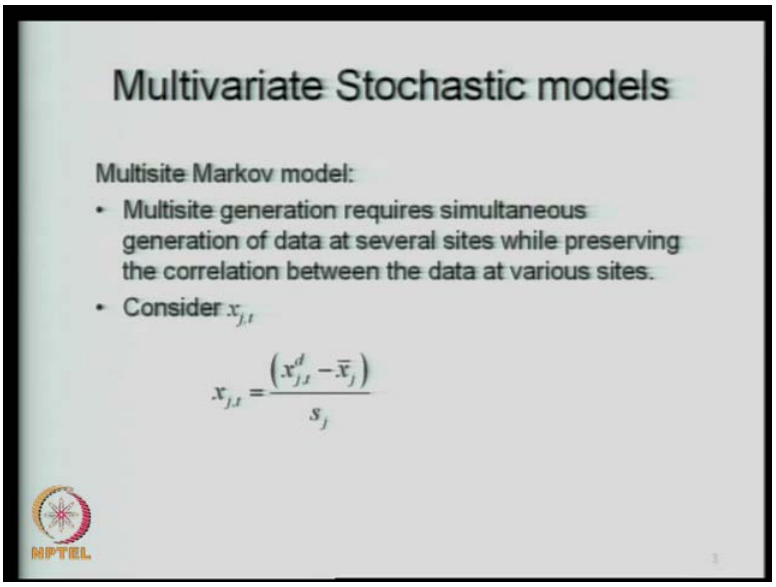
So, just to recall the expressions that we developed last time, first we introduced the concept of cross correlation at lag k. So, r j h is the cross correlation between the data at site j and that at site h, with a lag k. This lag k is time lag and this is how we determine this. So, this is at the i-eth time period and this is at i plus k-eth time period, where k is the time lag. This data refers to the site h data and this refers to site j data. So, this is how we get the cross correlation between data at two different locations and at two different times. So, there is a time lag of k time units between the two data that we are considering. So, i and i plus k. We also did the example, where we considered the lag 1 and then, how to found the cross correlation between two sites.

Then, we introduced the two site Markov model. This is for the subordinate site, where we are talking about generation at site h. You know, in these models, do not worry too much about this capitals and small letters. Actually, this should be a small letter, small x h t, but, like I did in regression, most of these the capitals and small letters, I mean the same thing, unless we are talking about the matrix notation, which will do subsequently.

So, we are generating the data at the site h for the time period t. We are using the data at site j, during time period t. So, this is data at site j during time period t. We are looking at the cross correlation between the data at site j and site h at lag 0. So, this is just the cross correlation between site j and site h at lag 0. The term u t here is the random term and we

have seen the expressions for these. This is again the cross correlation between the site j and site h at lag 0.

(Refer Slide Time: 05:22)



So, this is how we generate the data on the subordinate site, which incorporates the correlation between the sites, that is, the data at site j and site h. Let us see, how we can generalize this now to many sites. Now, from the two sites, we go on to many sites. When we are talking about a general multisite Markov model, the aim is to generate simultaneously at all locations. In the two site model, what we did was, in the two site model that I discussed in this particular course, what we did was that we identified a key site and then, generated the data on the key site, using single site model and on the subordinate site, using the correlations between these two sites, as well as autocorrelations. The term u t contains the autocorrelations there.

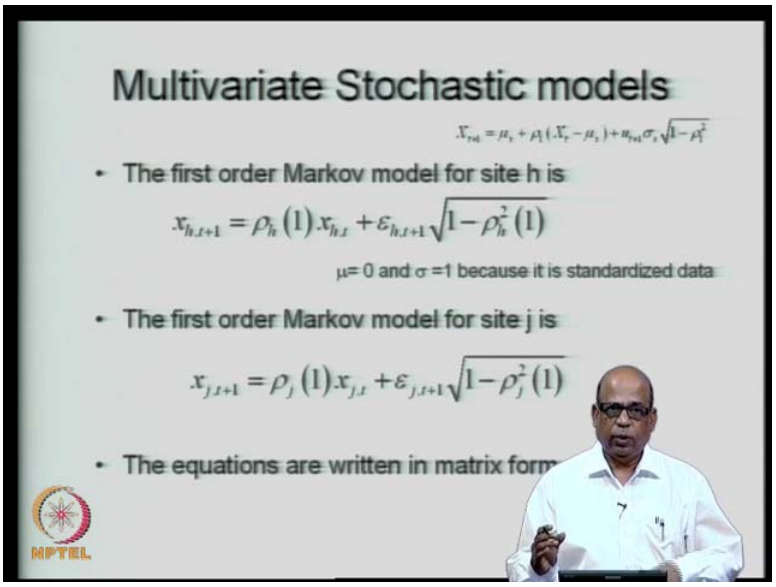
However, when we are interested in developing generation models for data generation, simultaneously at several locations, remember in the watershed, you may have three or four streams and your interest is to generate data on all these three or four streams simultaneously, taking into account their interdependence, as much as there is a correlation that exists among all these data.

When we are doing this, we do not go with the identification of key site and so on. We are interested in generating the data simultaneously at all the locations. So, there is no

site that distinguishes itself with other sites. All sites are equal in importance and we are considering the cross correlations among these data at various sites.

What we do is, first we work with standardized data. Let us say, this is the data now at the j-eth site for d, and the time period t. We remove the mean from this, deduct the mean from this and then, divide by the standard deviation of the data at the site j and then, form the standardized data. So, we will be dealing with only the standardized data x j t. Now, with this, we will start developing the multisite Marcov model.

(Refer Slide Time: 07:36)



Remember, we started with the single site Marcov model. What we now first do is that, we will apply the single site Markov model at all these locations. There are p locations, p stations. So that, you do not lose sight of the problem, we will just see what we are doing. There are several locations like this, 1, 2, 3 etcetera, p locations.

Now, all of these may be stream gages. For example, this may be a stream gage and this may be stream gage and this may be a stream gage and so on. So, we have the flow data at this location and we have the flow data at this location, etcetera. So, there are p locations at which you have the stream data. This may be a watershed. Then, you are interested in generating the data at all these locations simultaneously. That is the problem that we are addressing.

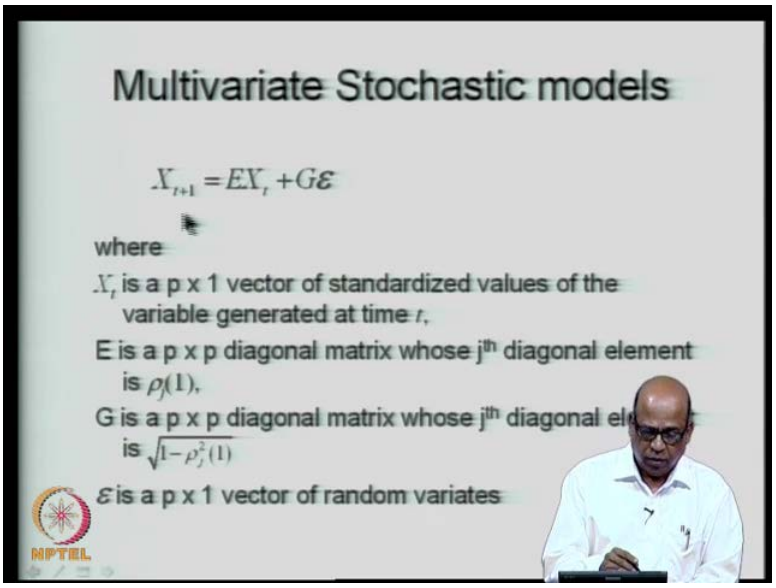
First, we take individual sites h j, etcetera. We take individual sites and then, we write the single site Markov model. Look at the single site Markov model. This is what we had developed earlier. This is a stationary model, in which the mean remains the same. Standard deviation and lag 1 correlation, they do not change from time period to time period.

So, this is a stationary Markov model. So, this is a single site Markov model. In this, because we have standardized the data, we can write this as rho h 1. We are talking about the site h for time period t plus 1. This is the generation model. So, for time period t plus 1, this is the lag 1 autocorrelation at site h. We are using the data for the previous period x h t because, we have standardized mu is 0 and standard deviation is 1. Because of this, now I will write, instead of u t plus 1, I will write e h epsilon h t plus 1 and this is 1 minus rho h square at lag 1. So, the rho 1, I am writing it as rho h 1, indicating that it is a lag 1 autocorrelation at site h. So, this is what we do at site h.

Similarly, for any site j, we will write the same expression for site j. So, like this for sites 1, 2, 3 etcetera, up to p, you can write these models. Now, these are single site Markov models applied to individual sites like this. When we combine these and then, want to write a single model for simultaneous data generation. The terms here, different terms that we are introducing for example, epsilon j t plus 1 that we are introducing, epsilon h t plus 1 etcetera, when we combine these, these terms must in some sense incorporate the dependence of the flows at site h, on flows at site j, for all h and j. So, that is what is the exercise that we will do now.

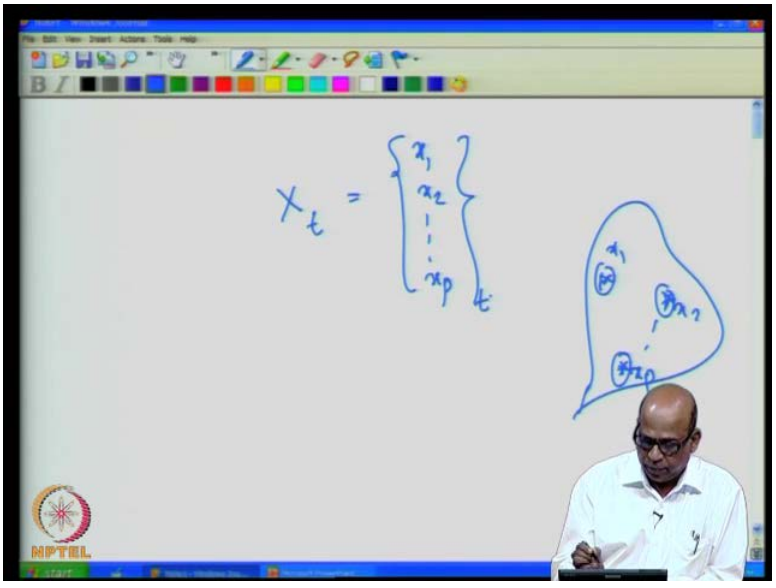
So, we write these expressions in matrix form and then, try to incorporate and then, in fact, incorporate the dependence of the flows at one location, on the flows at other location through their co variances and perhaps, correlations at different lags. In fact, we will consider lag 1.

(Refer Slide Time: 11:26)



So, we write this expression in a matrix form. We are generating x t plus 1 for p sites. So, x t plus 1 is a p by 1 vector of standardized values now. Then, we write this as e into x t. This is not expected value; e is a matrix, and e into x t. We are using for time period t plus 1. We are using the data at time period t plus g epsilon. All these are matrices.

(Refer Slide Time: 12:24)

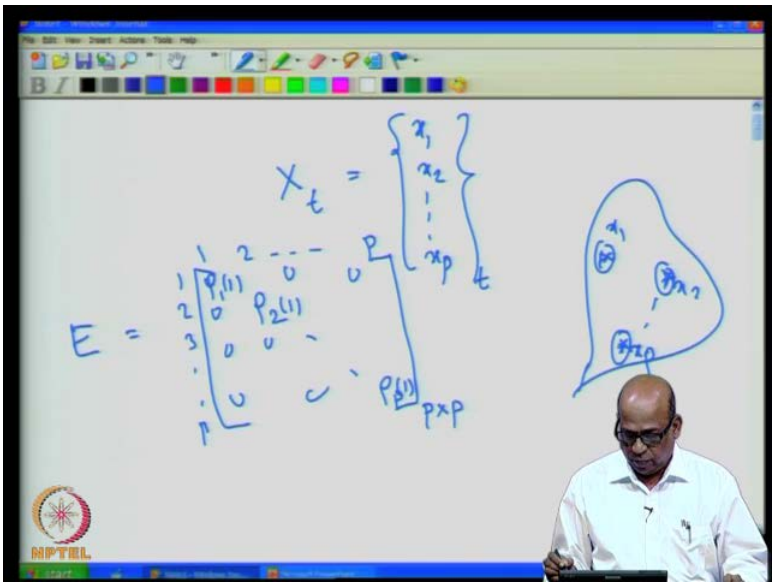


Now, x t is a p by 1 vector. This has standardized values of the variables at time t, which means let us understand all theses matrices correctly. So, we are writing this as, x t. x t is what we are considering at time period t and at each of the sites. So, this is x 1, x 2, x 3

excreta, x p. Now, these are the data at these various locations. Now, this is a thing and this may be x 1, this may be x 2, etcetera, x p; this is x p.

At each location, you have n time period data. So, you have x 1 at time period 1, x 1 at time period 2, etcetera. So, we are reckoning this at time period t. That is what forms x t. Then, we look at other matrix. e is a p by p diagonal matrix, whose j-eth diagonal element is, rho j 1. This is autocorrelation at the j-eth site at lag 1. So, let us write the e matrix now.

(Refer Slide Time: 13:32)

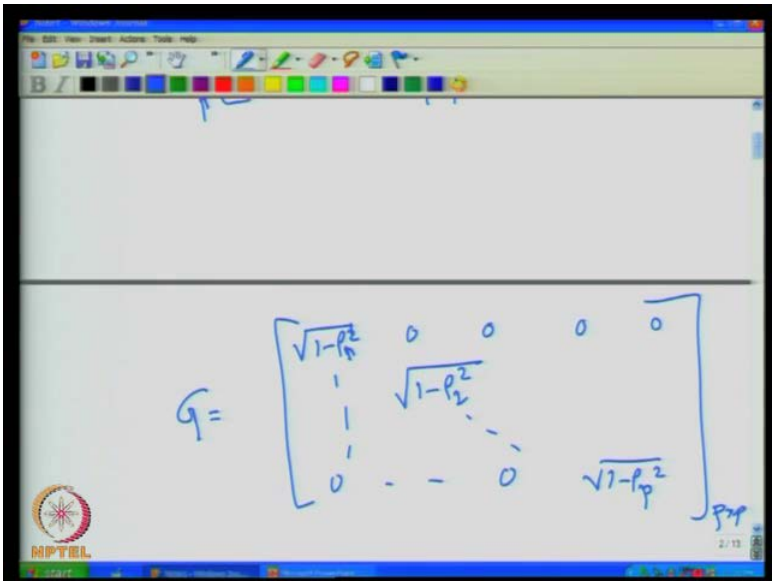


So, the e matrix can be written as, we have sites 1, 2, 3, etcetera, up to p. Therefore, you will have those many autocorrelations. You will have p by p matrix, 1, 2, 3, etcetera, up to p. It is a diagonal matrix, where only the diagonal elements are non zeros. So, I will write this as rho 1 and this is rho 2, etcetera, rho p and all of these are at lag 1. This is rho 2 at lag 1, rho p at lag 1 and so on. All other elements are zeros here. So, that is why it is called as a diagonal matrix. So, we form the e matrix by putting along the diagonals. The first, the lag 1 autocorrelations, this is 1 to 1, 2 to 2 and so on. This is how we get the p by p matrix.

All right. Now, we will formulate the other matrix now. So, e is the p by p diagonal matrix, whose j-eth diagonal element is rho j 1. That means, second diagonal is rho 2 1. That is, at site 2, the lag 1 correlation exists. That is what we put it as rho 2 1. Then, epsilon is a p by 1 vector of the random variates. Let us look at the matrix g first. Now, g

is a p by p diagonal matrix, whose j-eth diagonal element is root of 1 minus rho j square. Look at these. You know, these are the terms that we are trying to reproduce here. This is root of 1 minus rho h square 1. So, that is what we will reproduce through the matrix g.

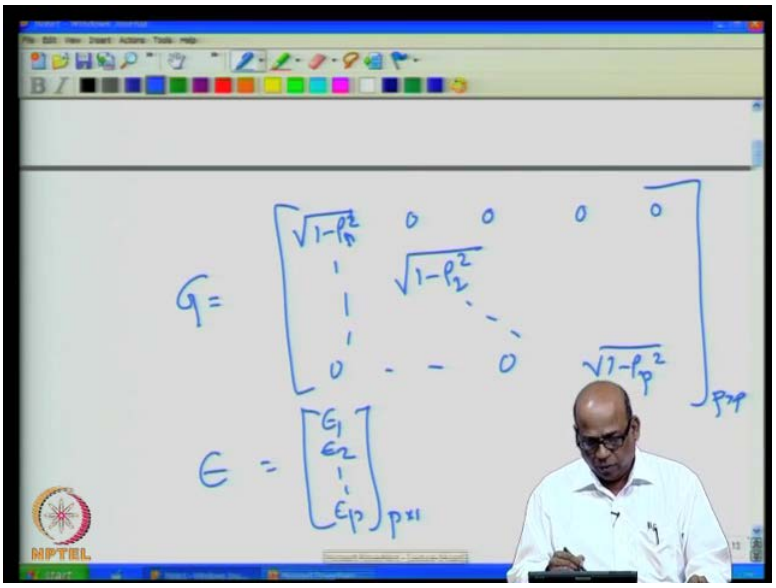
(Refer Slide Time: 16:12)



So, how does g look now? g will appear like this. I will write g as much similar to what we wrote for the matrix e. I will write g as root of 1 minus rho h square 1. That is, 1 indicates the lag and then, root of, you just look at the expression here. That is, p by p diagonal matrix, whose j-eth diagonal element is root of 1 minus rho j square 1. This indicates the site j and this indicates that it is a lag 1 autocorrelation that we are talking about. So, we will write this as, 1 and 1 minus rho 2 square and so on, root of 1 minus rho p square and all other elements are zeros. So, this is the diagonal matrix of size p by p.

Interpret the terms correctly here. All these are lag 1 correlations. rho 1, rho 2, rho 3, rho p, etcetera, they are all lag 1 correlations. Rho 1 here indicates the lag 1 correlation at site 1 and rho 2 here indicates a lag 1 correlations at site 2 and so on. This is rho p. So, this is how you formulate the matrix g. So, we got the matrix x, which is a standardized values and we have got the matrix e, which is a p by p diagonal matrix with the lag 1 correlations at that particular site. g is a p by p diagonal matrix, much similar to e matrix e, except that the diagonal elements are now, root of 1 minus rho j x square. Here, it was only rho j h rho j of 1 and epsilon is a p by 1 vector of random variates.

(Refer Slide Time: 18:08)



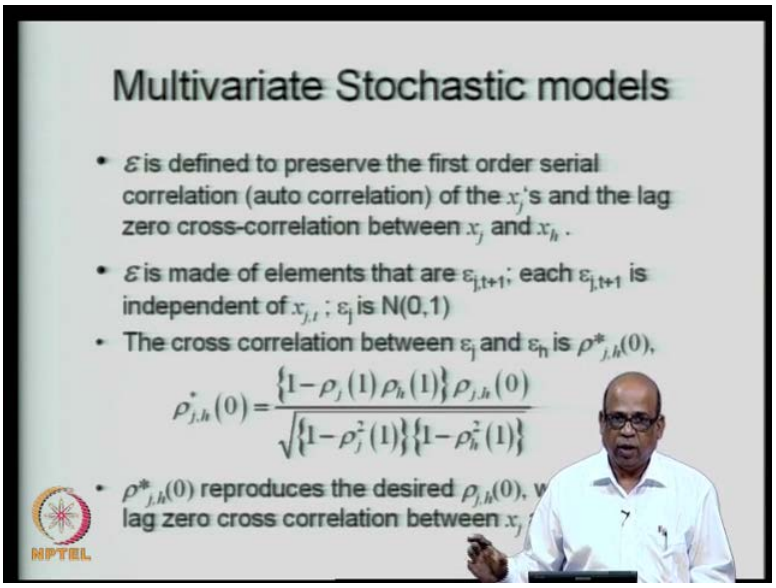
So, epsilon, we will write it as, let us say, I write epsilon. This is simply, epsilon 1, epsilon 2, etcetera, epsilon p. These are random variates. So, this is p by 1. So, this is how we write in terms of the matrix notation, the multisite model, that is x t plus 1 is equal to e x t plus g into e.

What are all known here? You look at all these matrices. x t is the vector of standardized data. This is known because, you have at p sites for the time period t, and you have the observed data. From the observed data, you have standardized. Therefore, x t is known.

Come to the matrix e. This is at site one. We are talking about the lag 1 correlation at site 1, that is rho 1 1. This can be estimated. That is, a time series available at site one and therefore, rho 1 can be estimated at site 1. Similarly, rho 2 can be estimated at site 2 and so on. So, all of these diagonal elements are estimated. Therefore, the matrix e is known.

Then, we will come to matrix g. Again, rho 1, rho 2, etcetera, are estimated. Therefore, matrix g is known. The epsilon, which is the random variates, they consist of e epsilon 1, epsilon 2, etcetera, up to epsilon p. So, if you look at this expression now, e is known, x t is known, g is known, and epsilon is what we have to define so that, your x t plus 1 gets generated. So, how do we define this epsilon? So, we are also looking at the cross correlations between different locations. So, epsilon, we need to define such that, the dependence of data at a particular location, on the data at some other location is captured by this model. That is what we will do now.

(Refer Slide Time: 20:52)



Now, epsilon is defined. The whole crux of this model now depends on how we define the epsilon. We define epsilon such that, we are able to preserve the first order serial correlation, which is the autocorrelation. We are talking about rho 1, rho 2, etcetera, all of them are at lag 1. It has to preserve the first order serial correlation and it should also preserve the lag 0 cross correlation between x j and x h. So one, the time dependence of the data at a particular site, which is given by the lag 1 autocorrelation and then, the space dependence based on the data dependency on other sites. That is reflected through cross correlations. So, you need to preserve the lag 1 autocorrelation, as well as lag 0 cross correlation. That is how we define epsilon now.

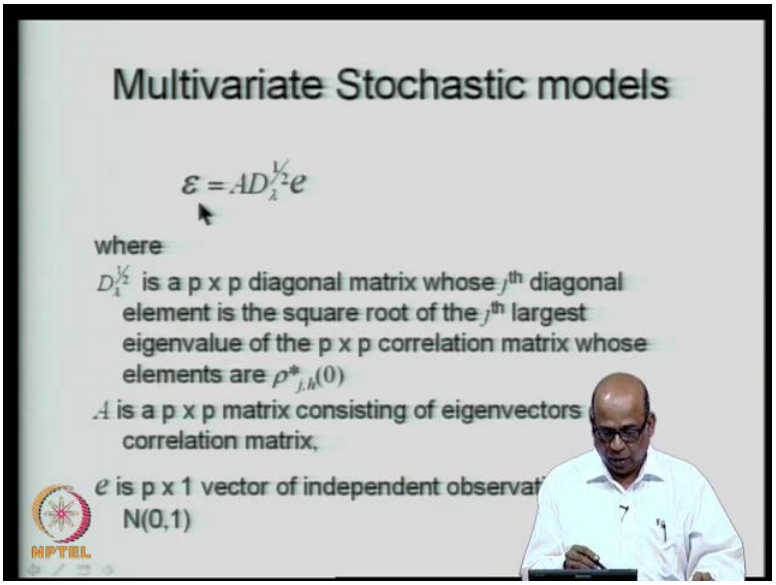
Now, we make the epsilon independent of x j t. That means, we have the original standardized data values. The epsilons that we are defining are independent of the x t's. That is one condition. The second condition is, we will use normal distribution. So, epsilon is also n 0 1, which means it has a 0 mean and standard deviation of 1. We define this rho j h star at 0 by this expression. Now, there are derivations etcetera, available. We will simply directly go to the application so that, we will be able to use these models in our applications. I will not go into the derivations of how we arrive at this.

So, rho j h 0. That is, this is the cross correlation between site j and site h at lag 0, at time lag 0 is defined by this expression. You look at what we have here. This is, rho j at lag 1, the autocorrelation at site j at lag 1, autocorrelation at site h at lag 1, and the cross

correlation between site j and site h at lag 0. Similarly, these are defined terms. We are defining another term here called rho j h star at 0, which depends on rho j h 0. So, this is the term that we define and we use this term to determine our epsilon.

If we define rho j h star 0 in this fashion, it reproduces, when we use this in our model as we will do presently, it reproduces the rho j h 0, which is the cross correlation between sites j and h at lag 0. Also the model that we have used, we have just introduced, it also reproduces the lag 1 autocorrelations. So, lag 0 cross correlations as well as lag 1 autocorrelations, is what is preserved by this model.

(Refer Slide Time: 24:20)



So, we define from this epsilon is equal to a d lambda to the power half e. It may appear to be mathematical and without any logic here. What we are doing now is that, we are obtaining the expressions for epsilon. What is the objective of this? We want to define epsilon, such that it preserves the cross correlation among different sites and it preserves the autocorrelations at a particular given site. So that, when we write the model in the form that we did and incorporate epsilon in this particular form, we are able to preserve the mean at all the locations, the standard deviations at all the locations, the lag 1 correlations at all the locations and the zero lag 0 cross correlations. So, this is what is needed for our model.

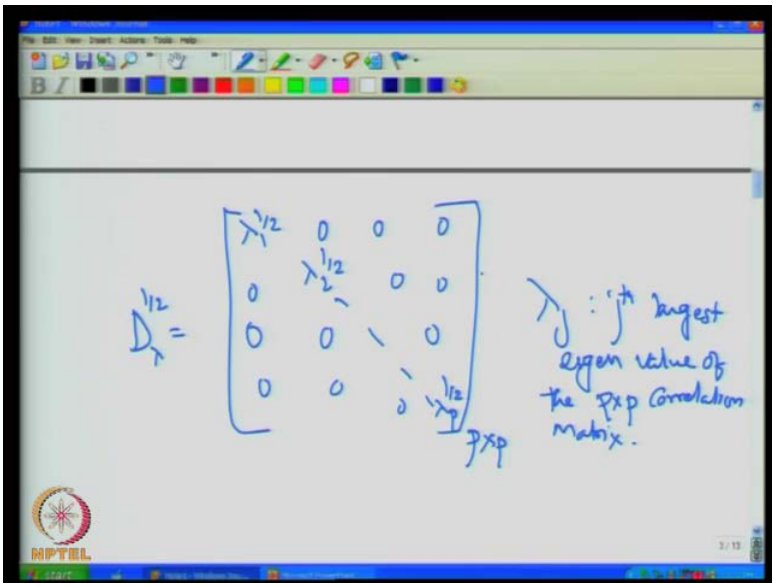
So, we define epsilon as a d lambda to the power half into e. We will understand what are all these now. d lambda is a p by p diagonal matrix, whose j-eth diagonal element is

the square root of the j-eth largest eigen value of the p by p correlation matrix, whose elements are rho j h star 0. Understand this correctly. This is what is necessary for our application. First you look at this correlation matrix defined by rho j h star 0. So, rho j h star 0 is what we had determined here. Between site j and h, you have rho j h star 0 defined.

On this now, you form a diagonal matrix. So, this is p by p correlation matrix. You have got by j h star. So, this is the p by p matrix. Let us write that. Then, we write the matrix d. So, for this p by p correlation matrix, which is a square matrix, you get the eigen values. We have seen how to get the eigen values for a square matrix. You get the eigen values for this square matrix and then, arrange the eigen values in decreasing order. The highest 1 as lambda 1 and next one as lambda 2, etcetera, up to lambda p. It is a p by p matrix and therefore, we have p lambda values, p eigen values. Arrange them in decreasing order.

Then, you look at the matrix d now. The matrix d will consist of lambda 1, lambda 2, lambda 3, etcetera, along the diagonal, where lambda 1 is the highest eigen value, lambda 2 is a next highest eigen value and so on, up to lambda p. That is how you get d matrix. But, we are interested in d to the power half, that is square root. So, you take square root of lambda 1, square root of lambda 2, and etcetera. So, that is how you form the d matrix. Let me write for clarity, what is the d matrix now.

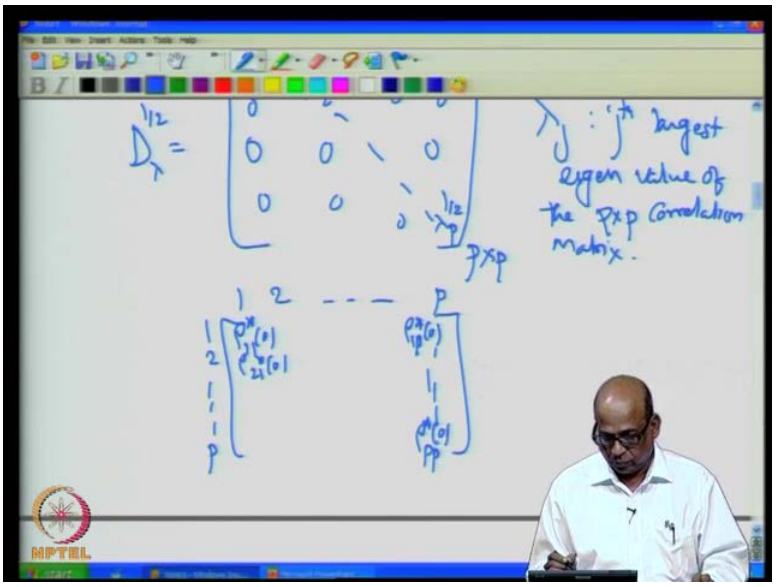
(Refer Slide Time: 29:12)



So, we write the d matrix as a diagonal matrix of size p by p. I say it is a diagonal matrix because, except the diagonal elements, all other elements are zeros. We write this as lambda 1 to the power half, lambda 2 to the power half and so on, up to lambda p to the power half. All other elements are zero and this is what I denote it as d lambda to the power half. All other elements are zero here.

What are these lambda 1, lambda 2 etcetera? I define lambda j as the j-eth largest eigen value of the p by p correlation matrix. So, this is how you form the d lambda to the power half matrix. I repeat again. You look at this. We are saying it is a p by p diagonal matrix, whose j-eth diagonal element is the square root of the j-eth largest eigen value of the p by p correlation matrix, whose elements are rho j h star 0. So, first look at p by p correlation matrix, which is rho j h 0, with each of its elements as rho j h 0. For clarity, may be, I will write that also.

(Refer Slide Time: 29:52)

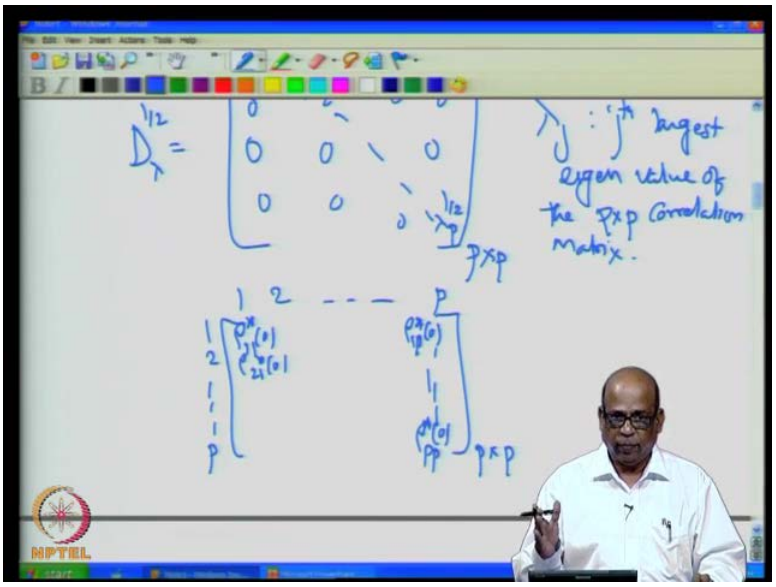


So, the correlation matrix that we are talking about is this, from which d is derived. Let us say, I have 1, 2, 3 etcetera, up to p and 1, 2, 3 etcetera up to p. Now, rho 1 1 rho j h star 0 and then, rho 2 1 star 0. Like this, any element you take. This will be rho 1 p star 0 and this will be rho p p star 0. That is the correlation matrix, whose elements are rho j h star 0.

Remember, these are not the original cross correlations. You look at this expression. This is what we will define rho j h star 0. So, at a site j with site h, this the cross correlation

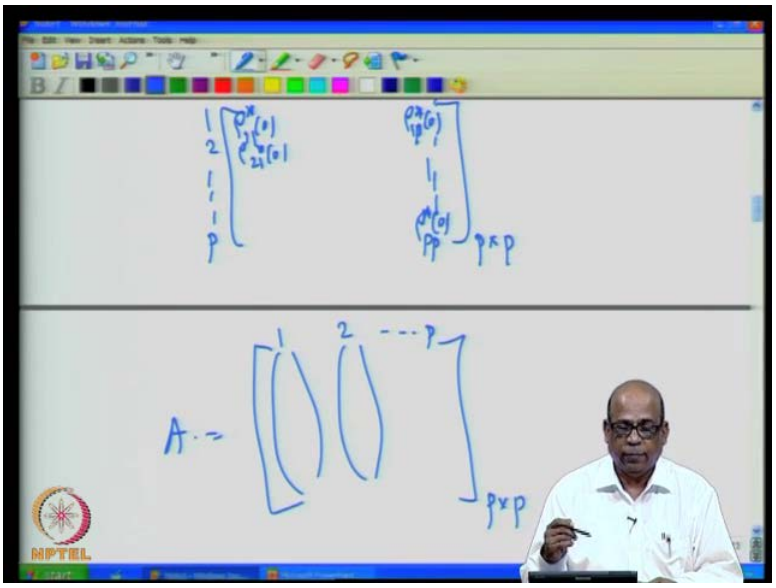
that you are defining and that is the cross correlation that you will use for defining the cross correlation matrix.

(Refer Slide Time: 30:58)



This is a p by p matrix. You take the eigen values. You will get p eigen values for this. Arrange the eigen values in decreasing order, with the highest one denoting as lambda 1 and the lowest as lambda p. Then, come to the formulation of d matrix. d matrix, is formed by the square root of, I mean, d lambda to the power half. That matrix, we form it as lambda 1 to the power half, lambda 2 the power half etcetera, as a diagonal matrix only. The diagonal elements are filled and all other remaining elements are zeros. This is how we formulate d lambda to the power half. All these values are known. Therefore, d lambda to the power half matrix is completely defined. All right.

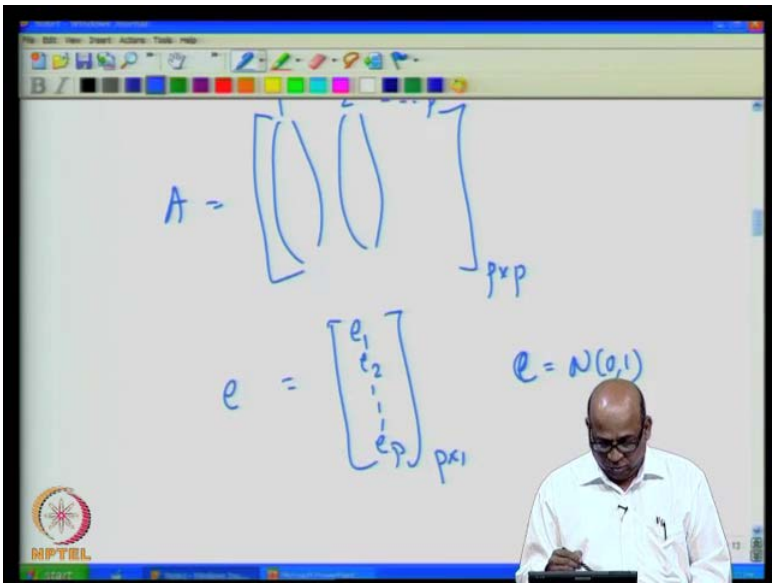
(Refer Slide Time: 32:36)



Then, we come to the matrix a. So, this is, we know how to determine d lambda to the power half. Let us look at what is matrix a now. a is a matrix, p by p size consisting of eigen vectors of the correlation matrix. So, we had the correlation matrix as we just defined. This is a correlation matrix and we were determining the eigen values. When we determine the eigen values, we also determine the eigen vectors of this matrix. The size is p by p and therefore, there are p eigen vectors. Then, we formulate the a matrix again as a p by p matrix. This one is eigen vector number 1, eigen vector number 2 etcetera. This is 1, 2, 3 etcetera, up to p. So, this is a p by p matrix. So, a is a matrix consisting of eigen vectors because, we are talking about eigen vectors of the matrix of size p by p. You have p eigen vectors. So, that is how you define the matrix a.

So, you defined the matrix d to the power half lambda associated with the eigen values and you define now, the matrix a based on the eigen vectors of the same matrix. This rho j h star 0 is a matrix that we are talking about. On this matrix, you define the eigen vectors and also define the eigen values. Thus, you define d to the power half lambda as well as the matrix a.

(Refer Slide Time: 34:12)



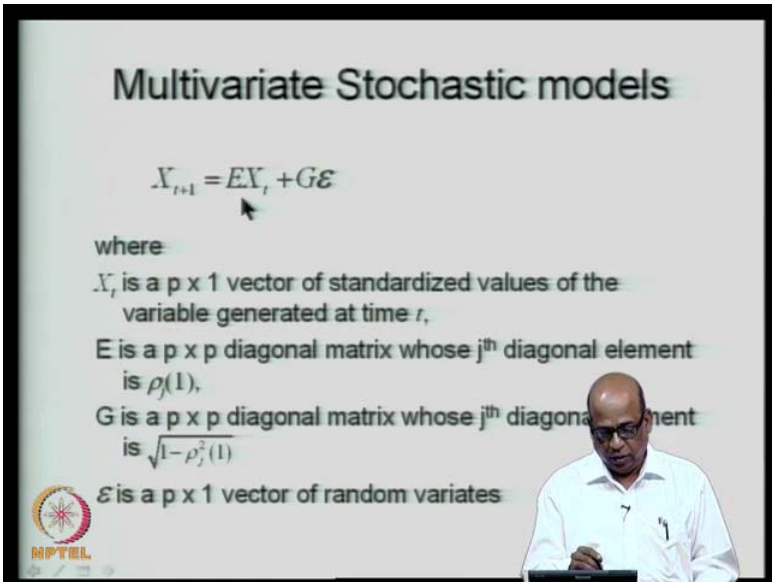
Now, we will come to e. What is this e? Now, e is the p by 1 vector of independent observations from n 0 1. So, this is essentially the standard normal variate. So, e is p by 1. So, I will write it as, may be, we will write it here only. So, e is, all right, we will write it on the separate sheet. So, e is now, I define this as e 1, e 2, e 3 etcetera, e p. So, this is the p by 1 matrix. What are all these? These are n 0 1. That is, standard normal deviates. So, you can just generate these from the normal distribution.

So, we have now defined all of these terms now. The matrix a is defined, matrix d lambda to the power half is defined and e to e is defined. Therefore, we should get epsilon. Now, a is a p by p matrix and this is a p by p matrix. So, this is p by p into p by p, and I get a p by p. Then, this is a p by 1 matrix. So, I get p by 1 matrix. So, epsilon is a p by 1 matrix. That is what we have defined here. Epsilon is a p by 1 vector of random variates.

What are all the information that has gone in to defining epsilon? We have essentially used rho j h star 0, which reproduces the way we have defined rho j h star 0. In fact, reproduces the cross correlations among different sites, j and h for all j and h. This is the information that we have used in defining epsilon. We use the eigen vectors of that matrix and we use the eigen values of that matrix etcetera, but, essentially we have used the elements rho j h star 0.

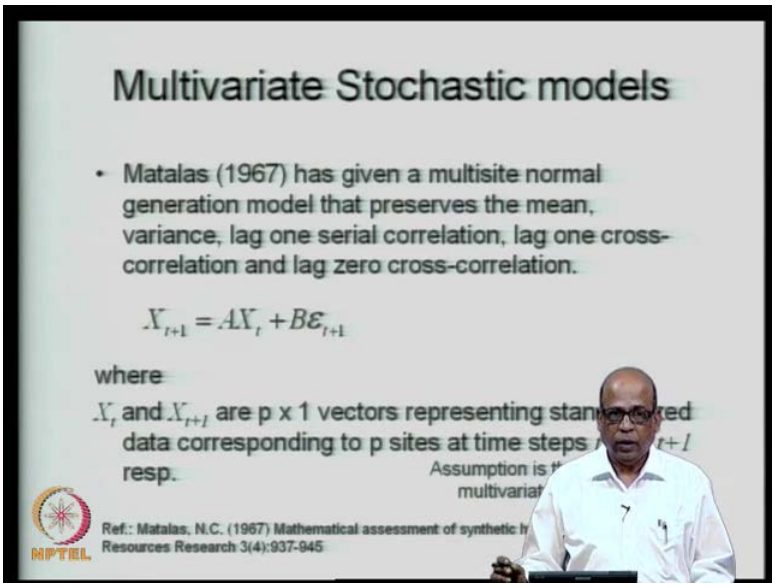
When we define epsilon by this, it essentially reproduces the dependence of values at one site on the values at several other sites. Remember, you are talking about j, site j and h can vary, and j h is equal to 1, h is equal to 2 etcetera. So, you are reproducing or you are capturing the dependence of the values at site j with values at several other sites, all other sites in fact.

(Refer Slide Time: 36:32)



Now, this is the general form. x t plus 1 is equal to e x t plus g into e. This is a multivariate stochastic model that we use in a most general form for simultaneous generation at several locations. This is a common general model. However, for specific case of flow generation, in several locations, in a watershed or in a catchment, a classic model is developed by Matalas. It is in fact called as a Matalas model way back in 1969. It is a very rigorous model. You know, it has been validated for several locations and we generally use for multisite data generation. We use the Matalas model in hydrology.

(Refer Slide Time: 38:12)



It follows the same broad structure, except that, what we did here is to incorporate our cross correlations among different sites. That information, we incorporated through epsilon. Here, what we did is, the lag 1 correlations, which means the time dependence, we have put through the vector e here or the matrix e here. This structure is slightly different in the Matalas model. Let us look at the Matalas model and how we use the Matalas model for multisite data generation simultaneously.

Following the same general structure of the model that we discussed just now, Matalas has proposed this model. This is a reference and this reference is, you know, is available in the C T Hann's book. I have taken it from the C T Hann's book. In fact, this area, classically it used to be called as synthetic hydrology. Thomas and Fehring, and Matalas, they have made a very significant contributions in data generation at several locations. In India, it has to pick up. These techniques have to pick up for water resource planning, management and operations etcetera, but, these techniques are available. They are very elegant techniques and they can be programmed easily. Therefore, the students of this course must be geared up to apply these kinds of stochastic models for actual applications, especially in terms of watershed management, in terms of the reservoir operation and so on.

All right. Now, look at this structure now. x t plus 1 is equal to a x t plus beta epsilon t plus 1 or we will call it as, b, the coefficient matrix. I call it as beta. This capital b also

indicates, b betas. Now, to generate x t plus 1, we are using x t and some random component here, epsilon t plus 1. Now, x t being the data, remember we have p sites. At each of the p sites, you have the time series t is equal to 1, t is equal to 2 etcetera. There are n time periods. So, x t is known and x t plus 1 is known. These we standardize. We use the standardized data. By standardization, I mean, deduct the mean and divide by the standard deviation. These are known.

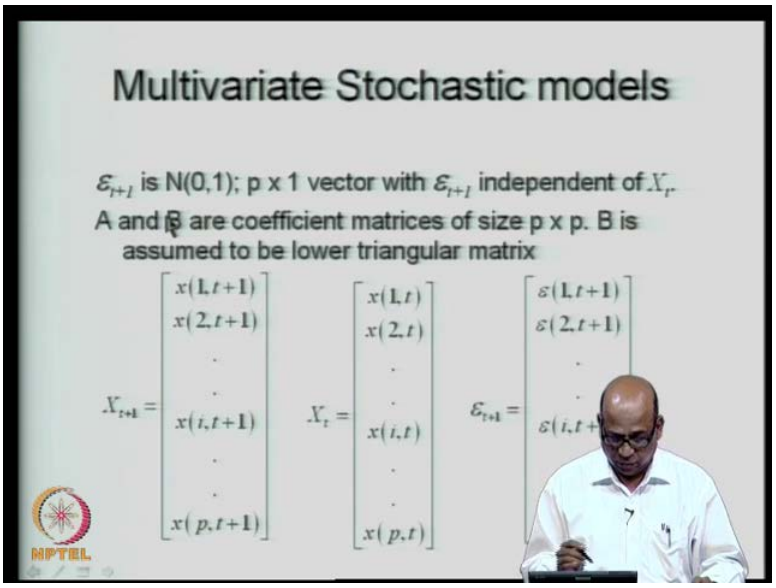
Now, we have to determine the matrix a as well as the matrix epsilon now. So, let us see how we do this. The Matalas model that we are just going to discuss, is a multisite normal generation model. It preserves the mean, the variance, lag 1 serial correlation and lag 1 cross correlation. In fact, lag 0 cross correlation also. So, it preserves all these parameters.

The mean, the variance, lag 1 serial correlation or the lag 1 autocorrelation at each of the sites and lag 1 cross correlation between different sites. You know how to determine the lag 1 cross correlation and then, the lag zero cross correlation. All of these parameters are preserved. What do I mean by preserved. When you do the data generation using this model, the generated data will have, in a statistical sense, same values for all of these parameters as the historical observed data. That is what we mean by preserving these parameters.

This is a normal distribution model. It uses normal distribution. The assumption is that, the model is a multivariate normal. What does it mean? At all the p sites, typically we are using it for flows. At all the p sites, the flows follow a normal distribution. That is what we mean by model is multivariate normal.

Now, these are p by 1 vectors. So, x 1, x 2, x 3 etcetera, x p, they are all normal. They all follow normal distribution is the assumption. So, this you should not forget. That is why we use these kinds of models for, let us say, monthly stream flow data, seasonal stream flow data etcetera. We cannot be using this for daily rainfall data, where obviously, the assumption of a normal distribution is not valid. So, whenever you are using such multi variates stochastic models, you must keep in mind that these are applicable, especially, the ones that I am discussing in this course are applicable only when individual variables x 1, x 2, x 3 etcetera, up to x p, they all follow normal distribution.

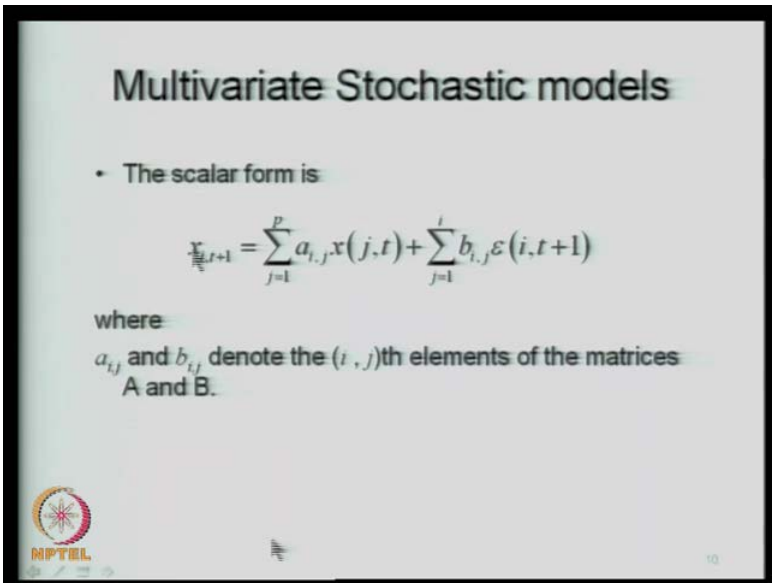
(Refer Slide Time: 42:44)



All right. Now, we will start defining what are these matrices now. Epsilon t plus 1, here this matrix is n 0 1. So, normal distribution, it is a standard normal distribution. It is a p by 1 vector, epsilon 1, epsilon 2 etcetera, up to epsilon p, at time period t plus 1. a and b are coefficient matrices of size p by p. Look at this. This is p by p, this is p by p and this is p by 1 and this is p by 1. So, you will get p by 1 here, for x t plus 1.

So, x t plus 1 is for time period t plus 1. I am writing x 1, x 2, x 3, x i and x p. So, these are the observations at different sites here, at time period t plus 1. Similarly, x t is observations at different sites at time period t. These are at different sites. There are p sites available and epsilon t plus 1, I will write as epsilon 1, epsilon 2 etcetera, at time period t plus 1. This is p by 1, this is p by 1 and this is p by 1.

(Refer Slide Time: 44:00)

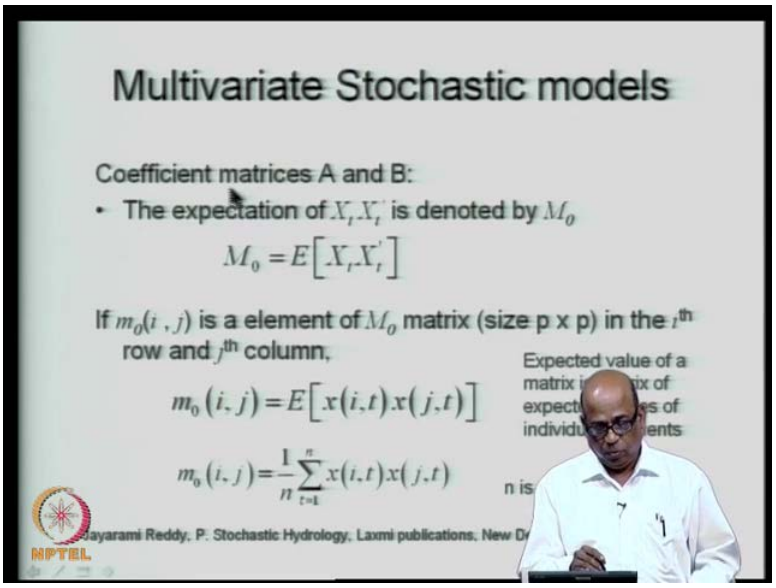


So, in the scalar form, if we write, that is this expression now, this expression I am writing in a scalar form. There is a matrix a and matrix b, which are both p by p matrices. So, I will write this as a i j. Look at this, I am writing for the i-eth site for time period t plus 1. The elements of the matrix a are a i j and the elements of the matrix b are b i j. So, I will write this as, j is equal to 1 to p a i j x j t and j is equal to 1 to i p i j epsilon i t plus 1. I am writing this. Now, there is a reason for going up to i. I will come to that.

We are assuming the matrix b to be a lower triangular matrix, which means all the upper values from the diagonal are all zeros. Only the lower values are not zeros. When we do that, we can write it in this fashion, that is, j is equal to 1 to i. I will explain this further, when we are developing the notations for matrix p. But, this expression arises out of the assumption that the matrix b is a lower triangular matrix. If we do not assume that, then we cannot find a unique solution. In fact, there will be infinitely many solutions possible. Remember, this is the kind of regression that we are doing, where you have to estimate the matrices a and b.

Look at this expression. a and b is what you have to estimate, because x t is based on the observations and epsilon t is based on your random numbers n 0 1. So, a and b are the ones that will determine these models. So, the exercise now is to define a and b. All right.

(Refer Slide Time: 46:36)



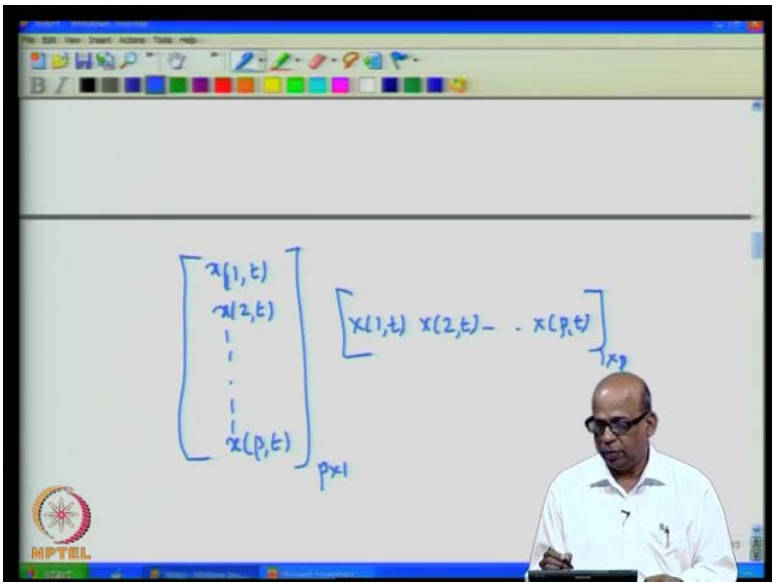
Now, this is a scalar form. I do not think it is necessary. As we progress, it will be clearer. Essentially, what we are doing is, we are using the elements of matrix a, a i j, elements of matrix b, which is b i j. We are assuming b to be a lower triangular matrix, in the sense that, all the elements to the right and top of the diagonal are all zeros. Now, the exercise is to determine the coefficients matrices a and b. Let us understand this correctly.

Let us say, we take the product x t x t dash. x t is a p by 1 matrix and this is x t dash will be 1 by p. So, you will get a p by p matrix. We take the expectation of this matrix x t x t dash, where x t dash is transpose of x t. We denote it by m naught m 0. Why m 0? We are taking x t and x t itself .That is the same time period. We can define m 1, which will, in fact, define m 1 as x t and x t minus 1 or x t plus 1, when we lag it by one time period. So, because we are considering the same time period, we will denote it as m naught and m naught is the expected value of x t x t dash.

In this now, can we write this x t and x t dash? Let us write this matrix, so that, it is clear. x t is a p by 1 matrix. You have x 1 t, x 2 t etcetera, up to x p t and x t dash is 1 by p matrix. x 1 t, x 2 t are arranged in a row. Then, we have to take a expected value. Remember, we are talking about expected value of a matrix. The expected value of a matrix is a matrix of expected values of individual elements. So, this is one thing that

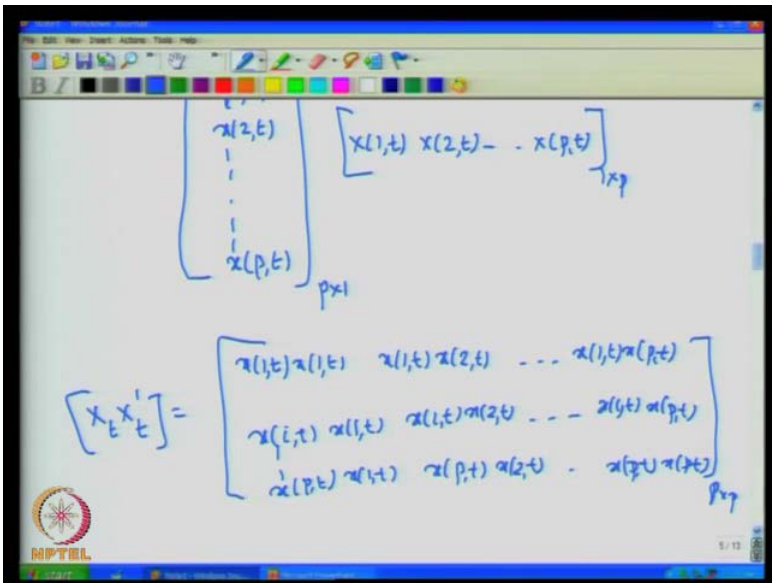
you must remember. So, let us see, how we get this x t x t dash matrix, because that is critical to the discussion that we will have further.

(Refer Slide Time: 48:46)



Your x t is a p by 1 matrix, consisting of the observations at p locations, standardized observations. So, I will write this as, x 1 t or to be more clear, we will write this way, that is x 2 t etcetera, x p t. Now, this is p by 1 matrix and this we call it as x t. This is x t matrix. x t prime, let us multiply it now directly. So, I will write this as, x t and I will write x t prime. x t prime is simply transpose of this. So, this will be x 1 t, x 2 t etcetera, x p t. This is 1 by p matrix. So, you should get a p by p matrix.

(Refer Slide Time: 50:24)



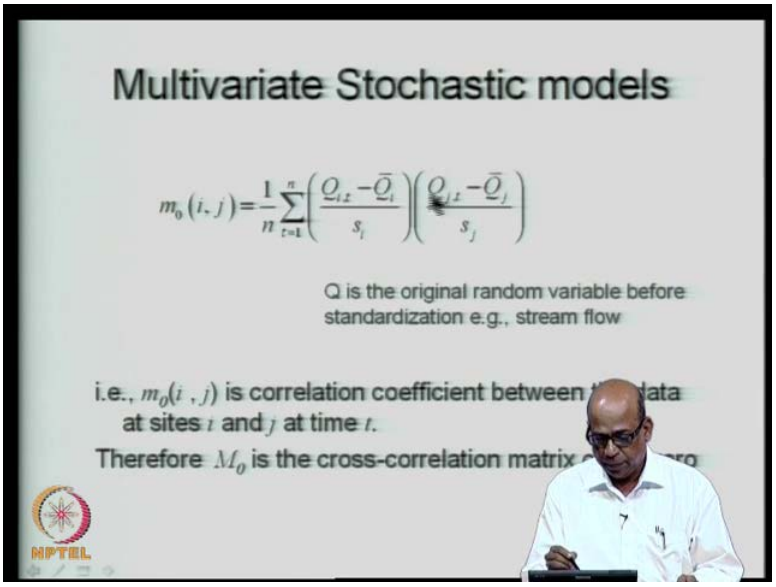
Now, that p by p matrix, we will write at least a few elements, so that, it will be useful in our further discussion. So, when I multiply this, what happens to the first row? Let us take this. The first row, I will define now. This will be x 1 t into x 1 t, x 2 t into x 2 t, x 1 t into x 1 t and x 1 t into x 2 t, x 1 t into x 3 t etcetera, x 1 t x p t. That will define your first row. So, let us write that x 1 t, x 1 t, this is a first element and then, x 1 t into x 2 t, x 1 t x 2 t etcetera, I come to x 1 t, x p t. This is how we get the first row. So, any i-eth row, if I want to write, it would be x i t, x 1 t, x i t, x 2 t etcetera, x i t x p t and last row is, x p t, x 1 t, x p t, x 2 t and so on, x p t, x p t.

What is it that I have done now? I have taken x t x t dash and that multiplication is what I have shown here. This matrix that we get, which is a p by p matrix, and this would be p by p and this is x t x t dash. So, this is what we got now. We have to take the expected value of this. So, we are talking about the expected value of x t x t dash. As I said, expected value of a matrix is a matrix of expected value. So, I will take expected value of this term, expected value of this term etcetera and then, form the matrix of expected values, which is the expected value of this matrix. That is what we will do now.

So, an individual element there, if you look at this point now, at any point, you have x i t, x j t. i j if you take or in this case, you are looking at an element x i t, x j t and then, you are taking the expected value of this. What do I mean by expected value? It is a time average that you are taking at that location, i t and j t and you are taking the time

average. So, when I take the time average, I can write the i j-eth element of the matrix m naught as 1 by n x i t, x j t. This is a estimate for the expected value. So, expected value x i t, x j t, I will simply take it as time average. This is x i t, x j t, and t is equal to 1 to n divided by n, where n is the number of time periods. That is how we define the matrix m naught, where all the elements are defined now and m naught is a p by p matrix.

(Refer Slide Time: 54:36)

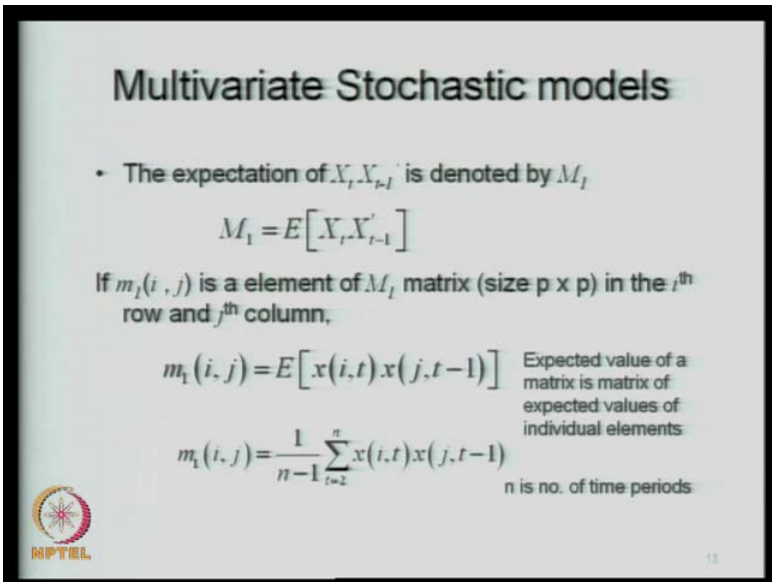


Now, look at this now, x i t and x j t. What are these? These are the standardized values of our original observations. Let us say, you are looking at the flows and we denote the flows as q. So, what is x i t? It is simply a standardized value of q, namely q minus q bar by s, where s is the standard deviation and q bar is the mean. Therefore, I will write this m naught i j, which is a i j-eth element in the matrix capital m naught as 1 by n, t is equal to 1 to n q i t minus q i bar. This is a original observed value and it is a mean and that is a standard deviation at site i and similarly, at site j. We are calling it as m naught because, we are reckoning it to the same time period t. t here and t here on both the sides i and j, where q is an original random variable before standardization. For example, it can be stream flow.

Can you look at this and identify what is this. This is q i t minus q i bar into q j t minus q j bar. So, this would have been the covariance, when you are summing it and dividing it by n and then, you are normalizing with s i and s j. Therefore, it in fact, indicates the cross correlation between the flow at site i and site j. So, m naught is, in fact, the cross-

correlation matrix at lag 0. We are not considering any lag here. We are taking about the time period t itself.

(Refer Slide Time: 55:48)



So, we know how to formulate the matrix m naught. We will progress and see how we formulate a matrix m 1 now, which will consider lag 1 with respect to time. So essentially, what we did today is that, we started with the multivariate stochastic model, where we are taking about preserving or taking into account the dependence of flows or dependence of the variable at one particular site on the variable at a different site. For example, you are talking about flows in a particular site to the flows in another site. We gave a general expression for these kind of multi variate stochastic models and then, we have just introduced the classical Matalas model, where we write x t plus 1 as a x t plus b epsilon t. Our purpose is now to estimate a as well as epsilon t.

The underlying assumption in this Matalas model is that, it is a multivariate normal model, which means that, all the variables have normal distribution. We are talking about site 1, site 2 etcetera, up to site p. You are looking at generation of data at all these sites simultaneously, by accounting for their dependence on each other.

So, we have just come up to, perhaps half way of developing the Matalas model. Our final aim is to arrive at this matrices a and b and also define the matrix, the vector epsilon, such that you will get x t plus 1. So, we are writing the expression for x t plus 1, which is the data at time period t plus 1, as dependent on data at time period t at all these

locations. This model preserves the mean, the standard deviation, the lag 1 autocorrelation, the lag 1 cross correlation, and as well as lag 0 cross correlation. All of these parameters are preserved by this model. Therefore, it is a powerful model in terms of data generation, especially for stream flows and especially for monthly stream flows, where your assumption of normal distribution is reasonably valid.

We will continue this discussion in the next lecture, where we will complete the development of the Matalas model and discuss an example to drive home the point of actual applications. Thank you very much for your attention.