Applied Multivariate Analysis

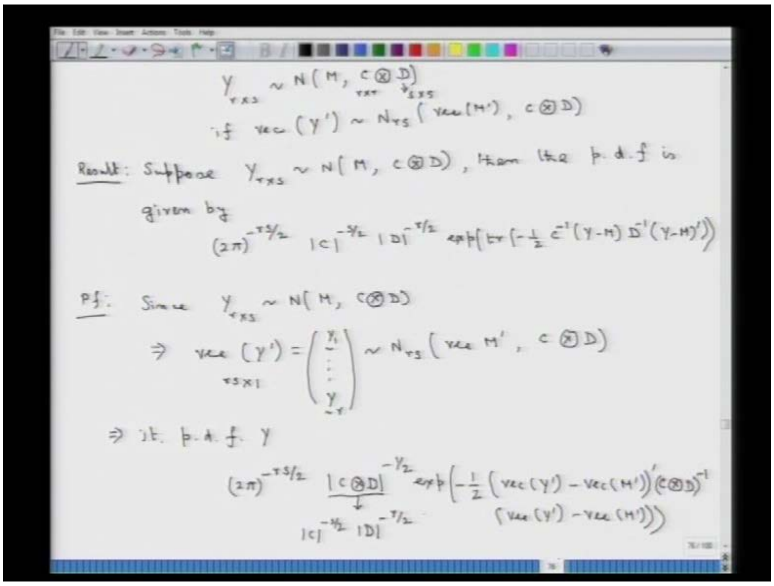
Prof. Amit Mitra
Prof. Sharmishtha Mitra
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture No. # 10

Random Sampling from Multivariate Normal Distribution and Wishart Distribution - III

Today we will first look at a few more properties about matrix normal distribution and also give an alternative definition of the Wishart distribution through a matrix normal distribution. And then we will be using these properties of matrix normal distribution it is probability density function and the alternative definition of the Wishart distribution in order to prove an important result in multivariate inference theory which is we will be proving independence of the X bar sample mean vector and sample variance covariance matrix that is S and also derived the distribution of these two important statistic to have X bar to have a multivariate normal distribution and S or rather n minus one times S to have a Wishart distribution.

(Refer Slide Time: 01:13)



So, let us first look at the following now recall that we had the matrix normal distribution defined in the following way that Y suppose is and r by s matrix random matrix. We said that we will call this Y random matrix to have a matrix normal distribution with parameters as one as m matrix and C chronica product D as the other set of parameters. If we have vec of Y prime to have a multivariate normal distribution r s dimension with mean vector as vec of M prime and covariance matrix as C chronica product D. Now, what is the p d f of this matrix normal distribution the p d f of this matrix normal distribution is nothing, but the joint p d f of the elements which are present in this random matrix .Now in order to derive this the joint p d f of the elements of this Y r s matrix we will use this particular result. That is the definition of a matrix normal distribution.

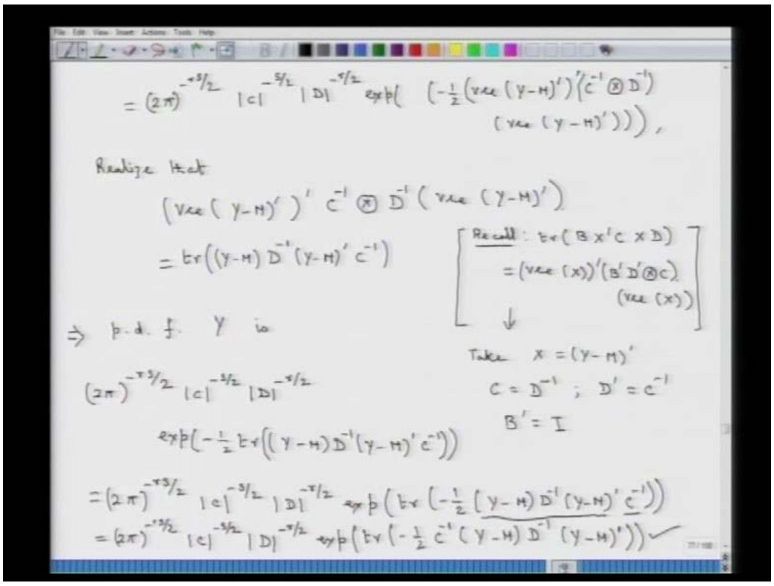
So, the result that we have is the following suppose Y r s dimensional random matrix has a matrix normal distribution M C chronica product D .Then the p d f the probability density function of all the elements of this random matrix is given by the following 2 pi to the power minus r s by 2 determinant of C. Now, in this recall that we have this dimensions that C is r by r and D is s by s. So, this is minus s by 2 determinant of D to the power minus r by 2, and then in the exponent we have e to the power trace minus half C inverse Y minus M D inverse Y minus M transpose. So, this is what is going to we are going to prove it that this is basically the probability density function of such a matrix normal distribution. Now how do we prove it?

Now since we have defined Y r Y r cross s random matrix to have a matrix normal distribution C chronica product D. So, this would imply that this vec of Y prime which we are going to have in stacks. So, this is say given by Y 1 Y 2 Y r. These are the constituent matrix vectors random vectors each of dimension s. So, this would follow multivariate normal N r dimension with mean vector as vec of M prime and the covariance matrix C chronica product D. So, if we have written it in this particular form we know what is the joint p d f of Y 1 Y 2 Y n, because that is the joint p d f of a multivariate normal distribution. So, this would imply that the joint p d f of Y basically we have this as a random matrix we have rearrange that in terms of this r s dimensional random vector.

So, the joint p d f of the elements of this Y matrix is same as that of the joint p d f of this vec of Y prime which is easy to write. That is the joint p d f of the elements here and that is basically is the p d f of a multivariate normal distribution with mean vector has vec of M prime and covariance matrix has C chronica product D. So, this is going to be given by 2 pi to the power minus r s by 2 then we have determinant of C chronica product D that is the covariance matrix. This to the power minus half then we have e to the power minus half then what is the vector here this is vec of Y prime this is vec of Y prime minus .It is mean vector that is vec of M prime and then the inverse of the covariance matrix. The covariance matrix is C chronica product D. The inverse of that multiplied by this vector itself.

So, its vec of Y prime minus vec of M prime this. So, this bracket n g M. This basically is the probability density function of the random matrix Y given through the multivariate normal distribution. Now, let us do some simplification to it.

(Refer Slide Time: 06:38)



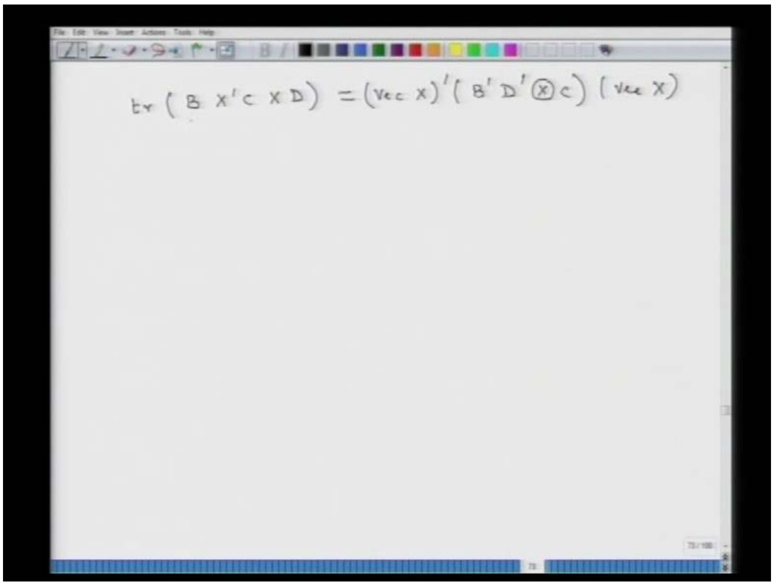
So, this is equal to 2 pi to the power minus r s by 2 .Now we had stated in the last lecture result concerning the determinant of this C chronica product D matrix. Now, this would be given by determinant of C multiplied by determinant of D raised to the powers determinant of C would be raise to the power of this D auto matrix that is s , and determinant of D would be raised to the power of the order of the C matrix that is r. So, we will be having that with the negative sign determinant of C. So, this term would lead us to determinant of C to the power minus s by 2 and this part actually. It is better just to write that what is this particular determinant. This determinant reduces to this multiplied by determinant of D to the power minus r by 2.

So, we can write that here. Determinant of C to the power minus s by 2 , determinant of D to the power minus r by 2 and then we have this exponent part which is e to the power trace minus half and .Then we have vec of Y minus M prime let me take this prime outside. So, Y minus M prime whole prime, because what we had here was vec of Y prime minus vector of M prime. One can write as vec of Y prime minus M prime , and then one can write that as vec of Y minus M whole prime and to the power prime and. There is this C chronica product D inverse. So, this once again we had stated what is going to be this C chronica product D is inverse. So, it would be C inverse multiplied by D inverse with a chronica product.

So, what will be having here is C inverse chronica product this D inverse. This is what remains here and then we will have this as vec of Y minus M prime. This is what is the joint p d f . Now let us concentrate on this exponent part and see what it reduces to realize that this part here vec of Y minus M prime whole prime C inverse chronica product D inverse into vec of Y minus M prime. Now, in order to simplify this we will use a result which we have once again stated. So, I will just keep it here recall this particular result that we had stated in the last lecture. We had stated that trace of a matrix B multiplied by X transpose C X D where B C X and D of course, confirm to this multiplication.

So, this would be equal to vec of X transpose then B transpose c transpose B transpose D transpose chronica product C this multiplied by vec of X. So, this is a result that we had stated sometime back I will just get you back to that particular result. This is that result number nine when we were discussing about elementary operations concerning this chronica product.

(Refer Slide Time: 10:43)



So, for this vec we had this result being stated. Trace of B X prime C X D was equal to vec of X transpose into B transpose D transpose chronica product C into vec of X. We are going to use this result now here. This is that particular result here. Now here in this result if we take the following values take our X the X that is sitting here to be Y minus M transpose which matches with this one we take C to be equal to D inverse. So, D inverse here is taken as the C here .Then let me take D prime which is sitting here to be equal to C inverse and B the matrix we do not see the presence of any third matrix here B.

So, we take this B prime to be an identity matrix. If we take that what does this reduced to. So, using this particular result and taking these special values has this X to be equal to Y minus M transpose as in here C has D inverse. Because that is what we will be requiring here the C which sits here is taken as D inverse in this expression and then D prime is taken as C inverse and B is taken B prime is taken as an identity matrix. See using that this would reduce to this is that term which matches with this one. So, that would be equal to trace of V is an identity matrix then X transpose what would be X transpose of this quantity. So, that would be Y minus M. So, that is X transpose and then C is what D inverse this would be equal to D inverse then we will have X.

That is Y minus M transpose that is X and then D and D is C inverse. We will have this as C inverse. So, this is what we are going to have if this vec is replaced by this particular trace. We can actually match this particular term there. So, this would imply that the p d f of this random matrix Y is given by 2 pi to the power minus r s by 2 .Let me just see it here.

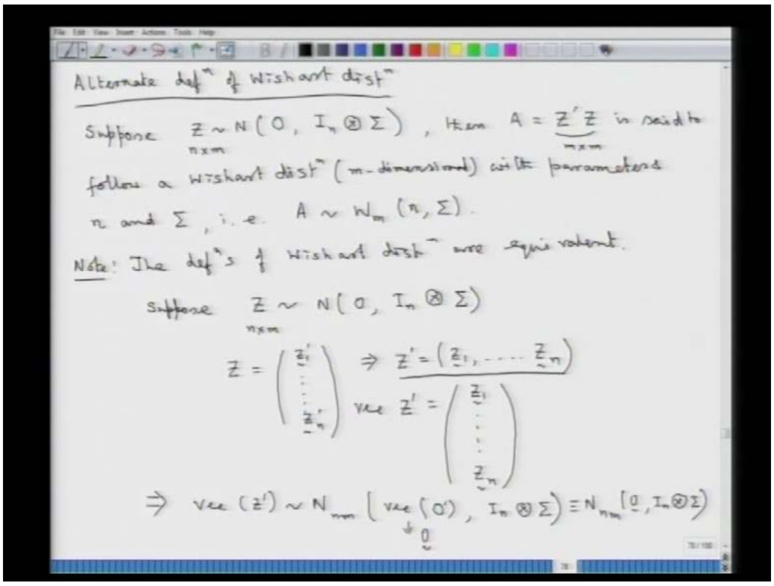
So, this is what we had there was no tress here actually. The tress will come later on as we see from the last page here what we had was the joint p d f of Y written in this particular form exponent to the power minus this particular term which is all in vec .Then this term in vec is what we have written here and that can be written as this particular trace. So, we replace that vec through this trace. We will have determinant of C to the power minus r by 2 determinant of D to the power minus this r by 2 this e to the power. Now this entire vec term here is going to get replace by this trace.

So, what do we have we have minus trace of Y minus M D inverse Y minus M transpose into C inverse this is trace sits here and this is sit. What we have is this as 2 pi to the power minus r s by 2 determinant of C to the power minus s by 2 determinant of D to the power minus r by 2 .Then exponent to the power trace minus half this trace of course, can be written in any way that we wish to write it. This D inverse Y minus M transpose C inverse and this is the form which we were supposed to prove only with this C inverse on the left hand side. So, what we can do is this is trace of this particular matrix. Take this to be one matrix and C inverse to be the other matrix. So, trace of a b will be equal to trace of b a and thus this finally, we can write in the form that it is required to be written. All those forms anyway are equivalent.

So, this is determinant of C to the power minus s by 2 determinant of D to the power minus r by 2 exponent to the power and then the form that we had stated there was stated minus half C inverse , because that comes here. So, trace of a b equal to trace of b a. So, C inverse comes in front. So, exponent to the power trace minus half C inverse Y minus M D inverse Y minus M transpose. So, this is the desired form which we were trying to prove derived. So, this was the joint p d f of the matrix normal distribution associated matrix normal distribution and this is what we have proved to be really having that particular form.

This form of the matrix normal distribution probability density function anyway is going to be used in the next important result which establishes the distribution of the variance covariance matrix or a constant multiplier of the variance covariance matrix and. The independence of the sample mean random vector with the sample variance covariance matrix.

(Refer Slide Time: 16:43)



But before we do that let me give this alternate definition of a Wishart distribution. Now, how do we define a Wishart distribution through a matrix normal distribution is what is giving us this alternate definition. Suppose we have Z follows a matrix normal distribution with say I take this Z the matrix normal distribution to be m by n say for example, in order to have tally it with the previous definition. Suppose I take this as n by m which is having a matrix normal distribution I it should be I n chronica product sigma. Now what is the order of this sigma. This sigma is what is going to corresponded to this particular order out here. So, this is going to be an m by n matrix.

Suppose we have a Z to follow a matrix normal distribution this then the quantity A which is a random matrix Z prime Z what is the order of Z prime Z. Z is n by m. This Z prime Z is a random matrix of order m by m .Then this Z prime Z is said to follow a Wishart distribution m- dimensional with parameters what are the parameters this is m by m. So, that is going to have the parameters as n and sigma. That is we will have this A to follow and m-dimensional Wishart with degrees of freedom as n and then associated variance covariance matrix as sigma.

Now, this definition looks a bit different than the definition that of the Wishart distribution that we had given in the last lecture .In the last lecture we had defined a

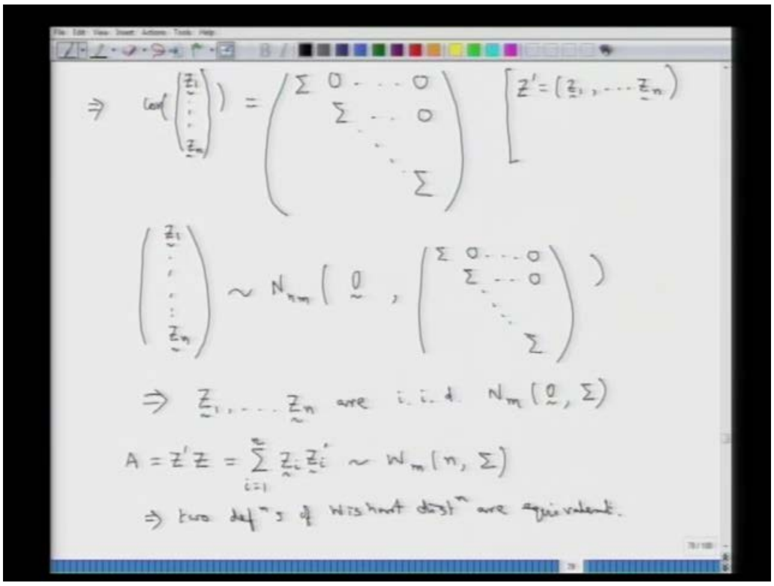
Wishart distribution through multivariate normal distributions. How had we defined. We had we had said that A is said to follow a Wishart distribution m- dimensional on degrees of freedom has n and sigma. If A can be written as summation of Y j Y j prime summation i equal to 1 to n where each of these Y js that is Y 1 Y 2 Y n are independent multivariate normal distributions with mean vector has null vector and a covariance matrix has sigma. That is how we had defined that Wishart distribution formation of a Wishart distribution.

This is what is giving an alternate definition of the same Wishart distribution through a matrix normal distribution. Here we have Z to have a matrix normal distribution with the associated mean matrix as a null matrix and the covariance matrix I n chronica product sigma of vec of Z prime. Then this quantity which is Z transpose Z is set to follow a Wishart distribution with parameters n and sigma. Now, we make an important note that the two definitions basically are equivalent .The two definitions of Wishart distribution are equivalent .Why are they equivalent .Suppose we start with this particular definition suppose we have this Z to follow we take this alternate definition and we will show that it also reduces to the first basic definition of the Wishart distribution.

So, this is what is the setup that is what we have. Now, this Z here is written as say Z 1 transpose Z 2 transpose and this is the nth row which is Z n transpose. So, each of these are 1 by m this is also 1 by m. This is leading this Z to be n by m. Since, this Z is said to follow A matrix normal distribution we will have from this Z defined Z transpose. So, this Z transpose is going to be Z 1 vector Z 2 vector and Z n vector.

Now, what is vec of Z prime from here vec of Z prime is nothing, but this Z 1 vector Z 2 vector all of them stacked one after the other. So, this is what is vec of Z prime .Now since it is given that this Z is having a matrix normal distribution with a null matrix here and I n chronica product sigma as a second set of parameter which will imply from this condition that Z is having this matrix normal distribution .This vec of Z prime will follow a multivariate normal distribution with dimension as m n. Just writing it as n m and with a mean vector as vec of this matrix here which is a null matrix this is going to give us a null vector of dimension n m and a covariance matrix as I n chronica product sigma. So, this vec of this null matrix is nothing, but a null vector itself. This is what we have that it is this vec of Z prime is having an n m dimensional multivariate normal with a mean vector as null vector and a covariance matrix as I n chronica product sigma matrix.

(Refer Slide Time: 23:12)



So, what does that tell us this implies that the joint distribution of this Y 1 Y 2 Y n is multivariate normal and the covariance matrix of this vec vector which is Y 1 Y 2 Y n this is going to be given by that I n chronica product sigma. So, I n chronica product sigma is going to generate this particular matrix .That it is a block diagonal matrix with sigma in all the blocks in all the diagonal in all the half diagonal blocks are this null matrices. So, what does that imply along with the fact that this joint distribution of vec of Z prime vec of z prime I had this Z 1 Z 2 Z n. I will just change these to Z 1 Z 2 Z n. That there is no confusion in the notations. This I had introduced as Z 1 this is Z 1 Z 2 Z n.

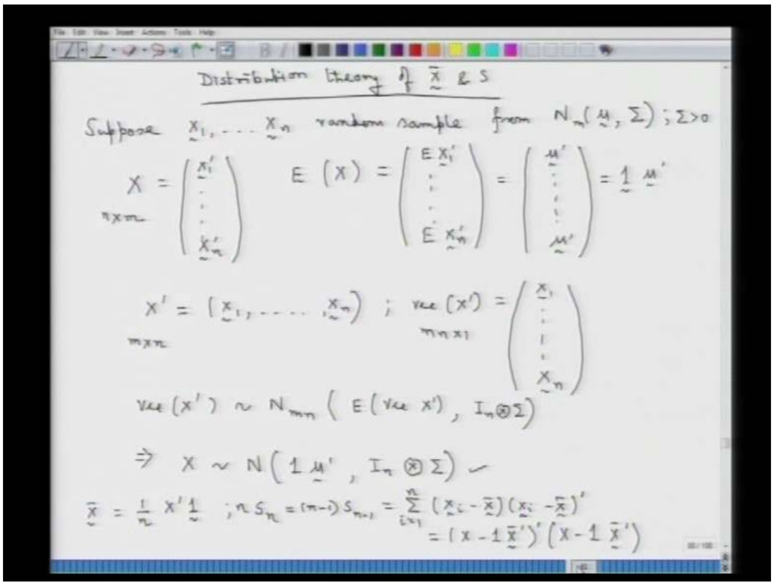
We have this that this Z 1 Z 2 Z n this random vector Z 1 Z 2 Z n this random vector follows a multivariate normal n m dimension with what parameters as the mean vector is the null vector and a covariance matrix as this. So, that is the covariance matrix sigma along all the blocks. So, this is that block diagonal matrix of diagonal blocks are all null matrices , this is what we have. So, this would imply that Z 1 Z 2 Z n each of them are going to be independent why, because the joint distribution is multivariate normal and the covariance matrix of Z i with Z j is all null matrices. So, this will imply that z 1 z 2 z n each of these are independent identically distributed ,multivariate normal m dimension with null vector as it is mean vector and sigma as its covariance matrix.

Now, this is the important pointer. So, what we have through the alternate definition of the Wishart is this A matrix which we had said is going to be Z prime Z .That is what the was the definition the alternate definition the Wishart distribution that if Z follows this. Then a Z prime Z is said to have Wishart distribution. Now, what is this Z prime. Z prime is given by this quantity which is Z 1 Z 2 Z n. What is the Z prime Z Z prime I will just write this Z prime here. Z prime was this Z 1 Z 2 Z n.

So, Z prime Z is nothing, but Z i Z i prime i equal to 1 to n simple. So, we have this particular quantity. Now realize what are these Z i is these Z i is Z 1 Z 2 Z n are i i d multivariate normal n dimension with null vector as it is mean vector and sigma as it is covariance matrix and. Hence this quantity a which is summation Z i Z i prime from the first definition that follows a Wishart distribution m n sigma. So, the two definition of Wishart distribution are equivalent. So, this implies that two definitions of Wishart distribution are equivalent.

Now ,we are in a position to go to the main result actually for which we have been actually looking at all these definitions of the Wishart distribution and alternate definitions multivariate matrix normal distributions and things like that .So, let us now move on to that.

(Refer Slide Time: 27:29)



Distribution theory of I say just of X bar and S. Now, suppose we have this we get back to random sampling from a normal distribution multivariate normal distribution .Suppose we have X 1 X 2 X n random sample from a multivariate normal distribution N p mu sigma where we assume that sigma is positive definite. So, there is no problem in dealing with that. Let me have this as m, because we have been using a notation m here. Now, from this X 1 X 2 X n if we look at the data matrix which now I pull in the following way. That this is X 1 transpose X n transpose.

So, what is the dimension of this is the data matrix. So, this data matrix has got the dimension that each of these now are 1 by m and there are n such rules. This is an n by m matrix. Now, what is the expectation of this n by m matrix this is that random matrix the random data matrix. So, that would be given by expectation of each of these vectors. So, expectation of X 1 prime expectation of X 2 prime and expectation of X n prime, what are these quantities each of them are mu primes. This is mu prime the second one also is mu prime and the last one also is mu prime.

So, I can write this as one vector which is having one on all the positions that multiplied by mu prime. So, this is what is going to give us this particular quantity here. Now, from this X the data matrix , what we are trying to do is to frame this random sample into a random matrix. So, that we will be able to say that what is the distribution of that random matrix in terms of a matrix normal distribution. So, the X transpose that is given or rather derived from this which is now m by n dimensional matrix which is going to be given by X 1 X 2 and X n. Now, each of these components here note that is a are coming from this random sampling from this multivariate normal distribution same multivariate normal distribution m dimensional with mean vector as mu and covariance matrix as sigma. So, this from here ,this would further lead us to this vec of X prime.

So, what is vec of X prime from here the vec of X prime from here is X 1 stacked over X 2 and so on X n. This is what is vec of X prime what is the dimension this is m n cross 1. Now, what is the characteristics of this vec of X prime. This vec of X prime is going to have a multivariate normal distribution. You can see that these X i components here in the vec of X prime they, this X 1 is independent of any of these. So, X 1 X 2 X n from the set of independent vectors here.

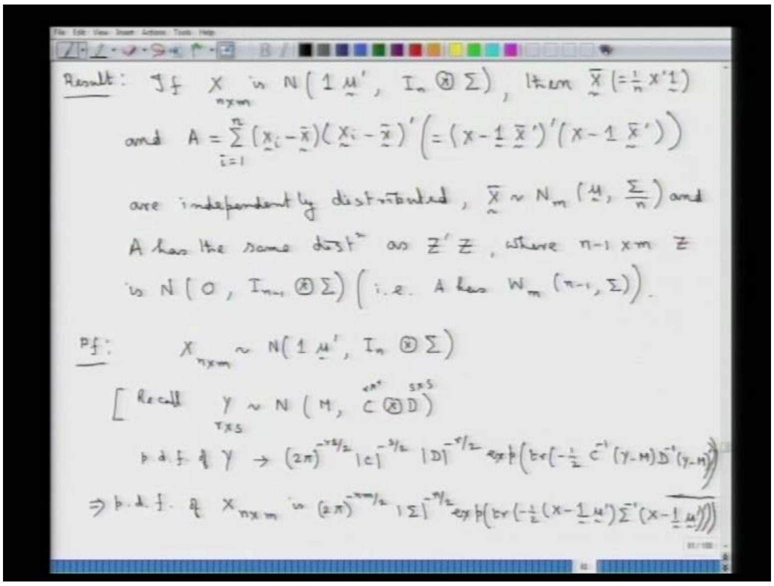
So, we will have this as n m dimensional multivariate normal distribution with what as mean vector .The mean vector would be given by this mu vector here mu vector here and that mu vector here. So, that would be this expectation of vec of X prime and what would be the covariance matrix. Now, since they are independent this would be given by I n a chronica product sigma. So, this would imply that this X the random matrix n by m will follow a matrix normal distribution with the mean matrix as this, because that is the expectation of this X whose vec actually coming out here. So, this is what we have as 1 mu prime as its mean matrix and the associated covariance matrix is I n chronica product sigma. So, this is what is going to play a major role that we have X the data matrix n by m which is formed from the random sampling from this multivariate normal distribution to have this matrix normal distribution.

Now, what are the quantities of interest that we are interested in finding the distributions this X bar and S. What are these two quantities is in terms of this data matrix. So, if we look at this X bar vector. This x bar vector is nothing, but 1 upon n X transpose this 1 vector ,why is that . So, if we look at this X transpose matrix here. So, it has got this is the first set of this is the first vector first random sample this is the second random sample and so on this is the nth random sample.

So, what we have here is that the first row of each of these vectors corresponding to the first observation. So, if we look at this X transpose i the first row which contains all the observations corresponding to the first variable gets multiplied with this one vector. It just gives the sum of all the observations corresponding to the first variable. The second entry in the X transpose one is going to give us the sum of all observations corresponding to the second variable and so on. So, we will have here in this X transpose one vector all the sum of the respective variables 1 to upto n and that divided by n is basically going to give us this sample mean random vector.

Similarly, one can have this n times S n or this is same as n minus 1 times X n minus 1. So, this is equal to as we have seen earlier this is summation i equal to 1 to n X i minus X bar into X i minus X bar transpose this we had seen. Now, this in terms of the X matrix the random matrix is X minus 1 X bar transpose whole transpose into X minus 1 X bar transpose this quantity. So, this n times S n or that is equal to n minus 1 times S n minus 1 which was given in terms of now, this in terms of the random vectors which we had from the random sampling and this is sample mean random vector. So, if we replace that by whatever we have got out here .We will actually be able to show that it reduces to this particular from this is trivial to show .Now, we will have this denoted by a matrix A and, then we will have this important result. I will first state this result.

(Refer Slide Time: 35:04)



And then prove it and explain what the what is the significants of this particular result and how this result is going to basically generalize the distribution theory of the univariate normal distribution. So, we have the result that if X which is n by m is a matrix normal distribution with the parameters as 1 mu prime and I n chronica product sigma. That is what is coming to us from this particular random sampling. So, this is what is the data that is what we have .Then this X bar random vector which we have already shown that to be 1 upon n X transpose 1 vector and, let us denote by a matrix A the quantity which is present in the sample variance covariance matrix without the multiplier either 1 upon n or 1 upon n minus 1 X i minus X bar into X i minus X bar transpose. We have shown that this A or rather we have not exactly shown we have stated , that this is equal to X minus 1 X bar transpose whole transpose into X minus 1 X bar transpose.

So, these two quantities if this is from a matrix normal distribution .Then X bar and A are independently distributed. they are independently distributed X bar follows a multivariate normal m-dimensional with a mean vector as mu and a covariance matrix as sigma by n and A has the same distribution as a Z prime Z ,where Z of course, is a matrix where n minus 1 cross m dimensional matrix . Z is having a matrix distribution with a null matrix as the mean matrix and I n minus 1 chronica product sigma as it is covariance matrix.

Now, in other words what we are saying is that A is having the same distribution as Z prime Z ,where Z the n minus 1 cross m-dimensional. So, this is what this has got a dimension which is m by m where each of(…) ,where this Z matrix n minus 1 cross m is having a matrix normal distribution this. From the alternate definition of the Wishart distribution thus we have if Z is having this matrix normal distribution. Then the distribution of Z prime Z would be a Wishart distribution that is A has a Wishart distribution m-dimensional with parameters has n minus 1 and sigma. So, the importance of this particular result is the following that this establishes actually. The independence of the sample mean vector X bar with the sample variance covariance matrix .Now, what is the sample variance covariance matrix ,the sample variance covariance matrix with the deviser n say is 1 upon n of this [ ] A matrix S n minus 1 is 1 upon n minus 1 of this A.

So, in this result we are saying that X bar is going to be independently distributed of A, that is the first part are independently distributed. So, this X bar quantity and this A matrix they are going to be independently distributed of one another. That is the sample vector is going to be independent of the sample variance covariance matrix and, furthermore it also gives us the distribution of X bar which independently. We have derived earlier .X bar is having a multivariate normal m- dimension with mean vector is mu and sigma by n as it is covariance matrix and. A has got a Wishart distribution m-dimension with degrees of freedom as n minus 1 and sigma as associated variance covariance matrix. So, this result perfectly generalizes the univariate distribution theory result .

Now, remember in univariate distribution theory when we had random sampling from a from an univariate normal distribution. We had X bar the sample mean univariate random variable in such a situation and X square or n minus 1 X square does not matter which we are looking at it is a constant multiplier. They were independently distributed X bar following a normal distribution and n minus 1 by X square was having a chi square distribution. This is what is the corresponding result in the multivariate distribution theory from random sampling from a normal multivariate normal distribution. Let us look at proving this important and fundamental result in multivariate distribution theory. So, we have this particular setup.

So, let us first start with this random matrix which is derived from the random sampling. So, we have this to have a matrix normal distribution I n chronica product sigma .Now we plan to look at what is the joint p d f ,we will start with the joint p d f of this X random matrix and then we will make a way make an orthogonal transformation such that. We will actually be able to associate one part of that orthogonal transformation with X bar and the other part of the orthogonal transformation with this Z matrix. Then we will show that this the two parts are independent and , hence the random variables which are random vectors or variables. Here random vector and random matrix derived from such parts of the orthogonal transformed random vector or random matrix will be independently distributed. In order to proceed we will require what is the joint p d f of this random matrix.

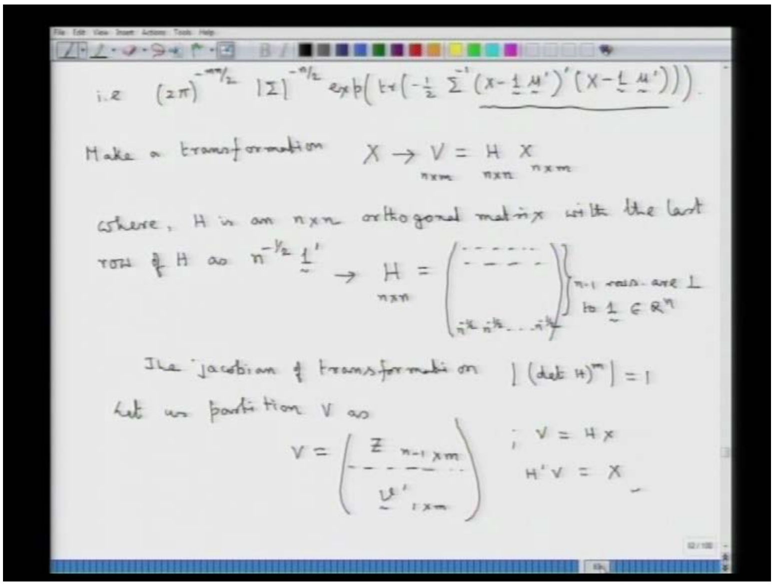
So, we recall this result which we had done today that if Y we had this result done today that. If Y has got a matrix normal distribution with mean matrix as M and C chronica product D as its covariance matrix .Then p d f of this random matrix Y was shown to be 2 pi to the power minus r s by 2 determinant of C remember this was r by r and this is s by s. So, this determinant of C to the power minus s by 2 determinant of D to the power minus r by 2 and exponent to the power trace minus half C inverse Y minus M Y minus M into D inverse Y minus M transpose. So, this is what we had here, where was it this was alternate definition and this is what we had derived the p d f of a matrix normal distribution. So, we are basically using this particular form out here.

So, we have this particular result. So, for this particular case here X is an n by m matrix normal distribution with this as m which is here, and this as a associated covariance matrix. So, the C in the general result will we will now take that to be I n the D in the general result we will take as sigma the m matrix is taken as 1 mu prime r is n and s is m. So, this would imply from this result. So, this is what we have recalled. So, this would imply that p d f of X this matrix normal distribution would be given by 2 pi to the power minus n m by 2.

Now, determinant of C .what is determinant of C, C is I n. So, determinant of C is one. This does not give any contribution. Now, our D in from the general result is sigma. So, what we will be having is determinant of sigma to the power of this C matrix r by 2. That is going to be minus n by 2 here, and then we will have in the exponent this trace of minus half .Now, what is C inverse, C is I n. C inverse also is an identity matrix and, then we will have X minus what is M. M is 1 mu transpose. That is the M matrix.

So, this is what is taking place of Y minus n then D inverse .Now D is sigma is our D. So, we will have this as sigma inverse and, then we will have this as X minus M matrix once again. So, that is 1 and then this mu transpose. This is what we have as the joint p d f of this X random matrix .It is actually useful to do something with this particular term which sitting in the exponent we will use trace of a equal to trace of b a and, then take this particular quantity this is X minus 1 mu transpose and. Then a transpose of that is sitting here, it is the transpose is here. There is no space here on the right hand side. So, we will have this is Y minus M transpose and then this is bracket for the trace and this is for the exponent.

(Refer Slide Time: 45:52)



So, let us take this in the form that would be useful. So, this is 2 pi to the power minus m n by 2 determinant of sigma to the power minus n by 2 and, then we will have exponent we will write it in the form that would be best suited for us. So, we will just write that as minus half sigma inverse now, comes on this side and what will be having is X minus 1 mu transpose. So, the transpose of this particular matrix was on the write and we have taken it to the left using the trace result and this is what we have remaining 1 mu transpose. This is what it makes this. Now, we make a transformation make a transformation from this X matrix to a matrix which is let us name that as V matrix which is equal to H times X this is an n by n matrix ,this is n by m matrix. So, What we have V also is an n by m matrix, where this H is an n by n orthogonal matrix with a special structure with the last row of this H matrix as n to the power minus half 1 transpose.

So, what is that we are saying here this basically tells us, that H this n by n matrix is having n rows. So, this is the first row here second row here and this is the last row here. So, this last row is n to the power minus half 1 transpose. All the entries in the last row are n to the power minus half. So, this is n to the power minus half. So, on this all these elements are n to the power minus half.

So, what does that imply that implies that each of these n minus 1 row which is sitting above the last row these are n minus 1 rows .Each of these n minus 1 rows are going to be orthogonal to this vector which is one. From the construction this n minus 1 rows are orthogonal to this 1 vector .Why is that so, because this H matrix is an orthogonal matrix. So, all the rows are orthogonal to one another. Now, since the last row is specified as n to the power minus half times one transpose. So, we will have every other n minus 1 rows of this H matrix to be orthogonal to this vector which is one, which will be useful in our present scenario.

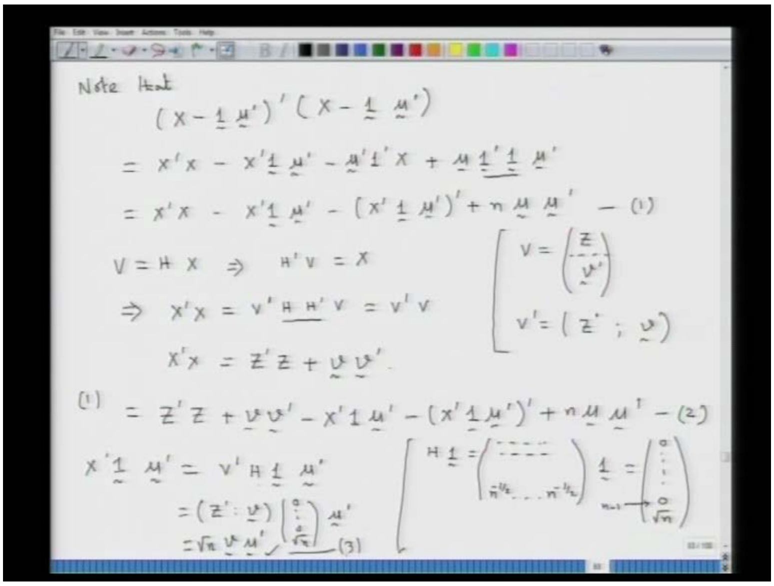
So, we have made this transformation from X to this V. Now, what is the jacobean of transformation .The jacobean of transformation jacobean of transformation from X to V would be given by the absolute value of determinant of H to the power m and that is equal to one, because H is an orthogonal matrix. Since, we have H to be orthogonal matrix this jacobean of transformation is going to be equal to one. So, what we are going to do is basically from the probability density function of the random matrix X. We are going to get into the probability density function of the random matrix V.

Now, let us write this V let us partition let us partition this V n by m matrix as following. Let us write this V as V remember is n by m. So, let us write that as Z which is n minus 1 by M matrix out here and the last will be a 1 by m vector. Let us write that as this V transpose where this V is a column matrix a column vector m by 1. We will have this as

1 by m the transpose of that. So, this is what we have as V. Now, note that from the transformation we have V equal to H times X.

Now, X is an orthogonal matrix. So, we will have this H transpose V to be equal to the X matrix, because H transpose H will be an identity matrix. So, we will have this in this particular form. Now, we will look at this exponent term here. Now, if we look at from getting to the joint p d f of this random matrix V from this random matrix X we will have to work with this work joint p d f which is the p d f of the ransom matrix X and then multiply that with the jacobean matrix which jacobean quantity which is nothing, but one in our present case and then replace this X is by the suitable quantities in the transformed random matrix. So, what we have to look at is to look at what is this quantity in terms of this transformed random matrix or its elements.

(Refer Slide Time: 51:46)



Now, let us work with that note that we have in the exponent X minus 1 mu transpose whole transpose X minus 1 mu transpose quantity. Now, let us look at what is this going to be equal to X transpose X this minus X transpose 1 mu transpose then ,minus mu 1 transpose X. So, this quantity is nothing, but just then transpose of this quantity this plus this we have a mu transpose I am sorry this is transpose of this quantity. So, it is mu 1 transpose 1 and then we have this as mu transpose. This is X transpose X minus X transpose 1 mu transpose minus X transpose 1 mu transpose whole transpose what is this equal to now, this term is 1 transpose 1. So, this is going to be just equal to n. So, this plus n times mu mu prime.

Now, what is this quantity and what is this quantity that is the point of interest .Now, X transpose X what is that. Now, we had this X being given by we had this V equal to H times X . Let me give a number to this , because we will be requiring that later stages. So, V equal to H of X. So, this implies that H transpose V this is equal to X. This would imply X transpose X will be equal to V transpose H H transpose V. So, this H H transpose will be equal to an identity matrix , because H is an orthogonal matrix. So, this is just equal to V transpose V.

Now, what is V transpose V in terms of the partition of that we had. We had introduced this partition that V equal to Z which was n minus 1 n minus 1 cross n n and this is what we had partitions partition as V transpose. From this partition if we look at what is V transpose V V transpose V is going to be Z transpose Z. So, that is Z transpose Z the transpose of this V multiplied with V itself and this plus V V transpose. So, this X transpose X is nothing, but this particular quantity. We have this which can be written in terms of this Z. So, one thus is equal to Z transpose Z these are this Z is matrix and this is V V transpose this is a V V transpose this minus we have not yet, addressed what is this part going to be this is write it in the form that it appears in the original expression this plus n times mu mu transpose .Let me give a number equation two.

Now, let us see what is this quantity equal to X transpose 1 mu transpose .Now, this is now, what is X transpose X transpose from here is V transpose H times one vector that multiplied by this is mu transpose here. So, we will have this as mu transpose here. What is this particular quantity. Now, remember what H is we had this H as this orthogonal matrix and, I had said that H has got special structure that its nth row is n to the power minus half and all the positions and .Hence that would imply that all the previous n minus 1 rows of H are orthogonal to this one vector one vector belonging to r to the power n. So, that is n dimensional vector.

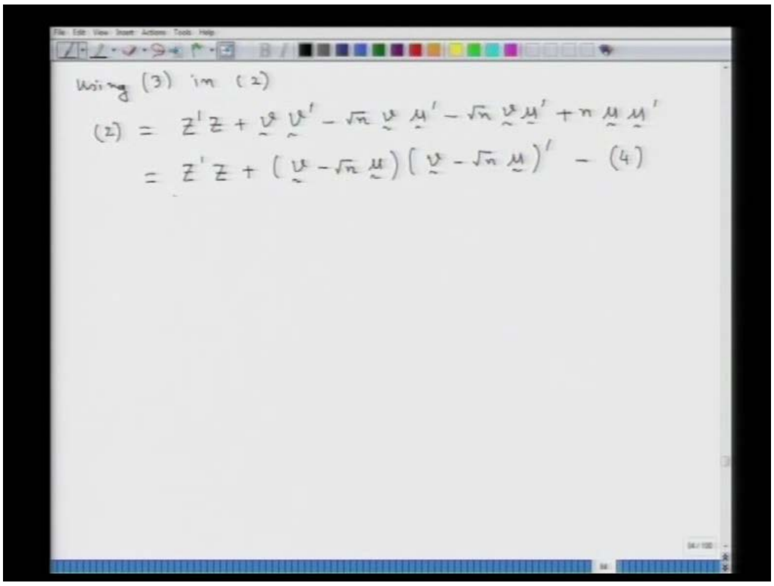
So, we have this particular special property of this H matrix. Using that special property what we can say is the following .Now, if we multiply H with 1 vector .Now, H is of the following nature that these are the rows of H first row ,second row and the last row is n to the power minus half at all the positions and if this is now, multiplied with 1 vector

.We had in the previous discussion said that the type of H that we are having n to the power minus half at all the positions at nth row .All the previous n minus 1 rows of H are orthogonal to 1.

And hence if we multiply H with one this with one would give us zero and. Will all the terms upto the n minus 1th row. All these terms are going to be 0 n minus up to n minus oneth position. This is the n minus 1th position of this H 1 vector and what is the last entry going to be this is going to be n to the power minus half now, this is n to the power minus half into 1 transpose. We will have a 1 transpose 1 which is n. So, we will have this as n to the power minus half into n and that is nothing, but root n. So, this H times i 1 actually ith H times 1 vector is nothing, but a vector n dimensional which is of this form that the first n minus 1 entries are zeros and the last entry is root n.

What do we have from here we have this V transpose as Z transpose that is augmented with this b vector, because that was what was the partition here. From here, we have this V transpose as Z transpose augmented with this V vector. That is the form of V transpose and then H times 1 is nothing, but our zero on all the positions up to n minus 1 and the nth position is root n and that multiplied by this mu transpose. So, what is this equal to this is simple because, the first n minus 1 entries are zeros and thus this just is equal to root n times V mu transpose. So, this is simple that this X transpose 1 mu transpose is just equal to n to the power half V mu transpose. We will use this particular form in expression number two.

(Refer Slide Time: 58:55)



So, let us do that. Using let me give an equation number here. This say is equation number three using three into[ ] using three in expression number two what we have is this Z transpose Z that expression number two is equal to Z transpose Z this plus V V transpose and then all the entries using that expression three we will have this as root n V mu transpose. The next entry is exactly the same root n V mu transpose this plus n times mu mu transpose and this can be written as let me keep it as it is Z transpose Z. This plus if we look carefully here this is V minus we can write it as root n mu into V minus root n mu transpose.

This is what is finally, the exponent what we had here was this particular term which was actually the term involving X quantities and sitting in the exponent .Now, that in terms of the transformed random matrix takes this particular form. So, we will end this lecture at this particular point and then in the next lecture we will look at this form equation number four and. Then we will look at proving the result that we have stated which establishes the independent of X bar random vector and S the sample variance covariance matrix .thank you.