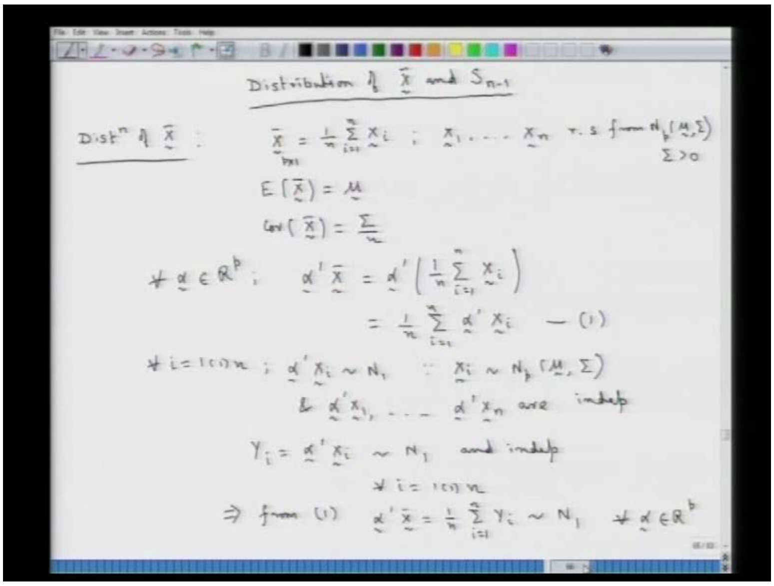
Applied Multivariate Analysis

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

Lecture No. # 09
Random Sampling from Multivariate Normal Distribution and Wishart Distribution-II

In the last lecture what we had actually seen is we had X bar the sample mean vector to be an unbiased estimator of the population mean vector mu. The covariance matrix is the sample variance covariance matrix either S n or S n minus 1. Both of them are a sufficient all though they are sufficient statistic, but only one of them S n minus 1 that is the sample variance covariance matrix with a deviser n minus 1 there is the unbiased estimator of sigma the population variance covariance matrix X bar, of course is a unbiased estimator and also the sufficient statistic corresponding to mu only under the situation that sigma is known. We had also proved that X bar and S is jointly sufficient for mu and sigma and the other cases when say mu is known to be mu not and all we had seen that. We are also derive the maximum likelihood estimator of the respective parameters associated with the multivariate normal distribution say, X bar the sample mean vector was the maximum likely would estimator of the population mean vector mu and S n the sample variance covariance matrix with a deviser n was shown to be the maximum likelihood estimator of the population variance covariance matrix sigma under the multivariate normality set up. What we will look at today is basically the distribution of these statistics that we had discussed in the last lecture.

(Refer Slide Time: 01:58)



So, we will be looking at the distribution of the associated statistic in this lecture. Distribution of X bar the sample mean random variable vector and S with a say deviser n minus 1 .The distribution of S n would also follow in a similar manner. First let us concentrate on looking at what is the distribution of this X bar random vector. Now, X bar random vector it is based on a random sample of size n. So, that is given by 1 upon n summation X i, this summation is from 1 to upto n. Now, what are these X i components X i components are basically the n random samples that are observed.

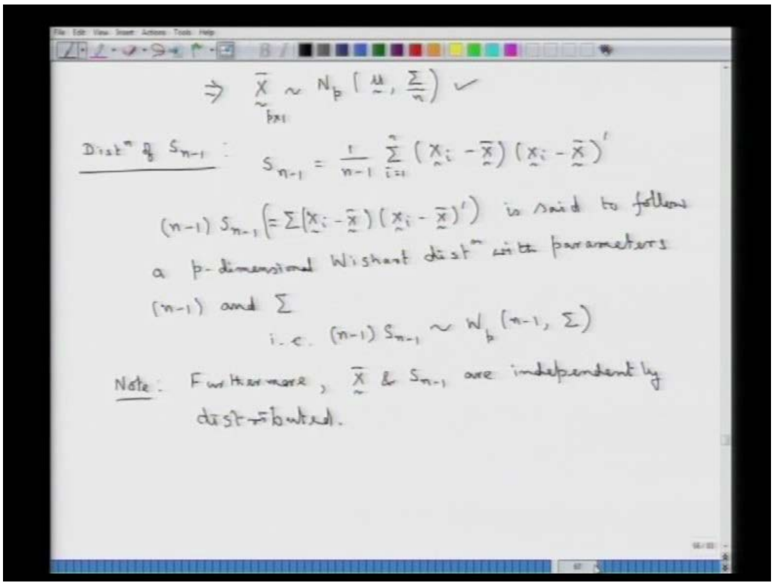
So, this X 1, X 2, X n is a random sample from multivariate normal distribution p dimensional with a mean vector mu and covariance matrix as sigma, sigma is positive definite. Under such a situation what is the distribution of this X bar the sample mean random vector. See what we have already obtained is the following that Expectation of X bar is mu that is this is unbiased estimator of mu and. we have and also derived the following that the covariance matrix of this X bar the sample mean random vector irrespective, of course multivariate normality on the underline distribution is sigma by n. Now, in order to see what exactly is the distribution of this X bar random vector we will look at the definition of multivariate normal distribution. This is p dimensional vector here.

So, for every alpha vector belonging to R to the power p. Let us see what is the distribution of this alpha prime X bar and what can we say about this alpha prime X bar quantity. This alpha prime X bar is nothing, but alpha prime 1 upon n summation i equal to 1 to n X i quantities. So, we can write this as, 1 upon n outside and inside the summation what we have is this summation i equal to 1 to n alpha prime X i. Now, each of these X i inside this summation which contains n terms are multivariate normal. So, X 1 X 2 X n all of them being a random sample from this N p mu sigma has the following properties that, each of this are multivariate normal N p mu sigma and they are independent.

So, what we can say is that for every i equal to 1 to upto n .This alpha prime X i linear combination is going to follow and N 1 distribution. Why this is , because each of this X i follow a multivariate normal p- dimension with mean vector mu and covariance matrix sigma. Hence, this basically are in one random variables and further more ,because X 1 X 2 X n is a random sample these X 1 X 2 X n are independent and. this alpha prime X 1 alpha prime X n are independent. This is because X 1 X 2 X n are independent. So, if we denote this Y i equal to alpha prime X i , these are following N 1 distribution and all of them are independent they are actually identical also. It is basically an i i d set up this is true for every i equal to 1 to upto n. So, what does that imply if we have these alpha prime X i which we have denoted by Y i. So, this term alpha prime X bar is nothing, but 1 upon n summation i equal to 1 to n Y i where each of these Y i s are N 1 independently distributed. This is what this is just the some of n independent normal distributions. So, this would imply let us give this equation number one.

This implies from one and the discussion that we had about this each of these alpha prime X i s. This would imply that this alpha prime X bar which is 1 upon n summation of this Y i quantities, Each of them are normally distributed random variables and Y 1 Y 2 Y n are independent. So, this would imply that this is going to have an N 1 distribution and univariate normal distribution this is going to be true for every alpha, belonging to R to the power p where of course, alpha is not a null vector. What were what have we proved we are proved that for every alpha belonging to R to the power p this alpha prime X bar has got N 1 distribution. So, that would imply that the distribution of X bar is multivariate normal, because that is what is the definition of multivariate normal distribution.

(Refer Slide Time: 07:43)



This would imply that this X bar the sample mean random vector this follows a multivariate normal p dimension with a mean vector as mu and covariance matrix as sigma by n. So, this is the desired distribution of the sample mean random vector which is p-dimension derived from the set of random samples X 1 X 2 X n. This is relatively simple, what is new actually is the distribution of S n minus 1 say .Now S n minus 1 is given by 1 upon n minus 1 summation i equal to 1 to n X i minus X bar into X i minus X bar transpose .Now, this S is say to have a Wishart distribution which is new distribution of course, we are going to talk about what is Wishart distribution its properties and how this Wishart distribution is going to be derived. Let me first just state the result, this n minus 1 S n minus 1 this is summation X i minus X bar into X i minus X bar transpose is said to follow a p this is dimension is p that is right.

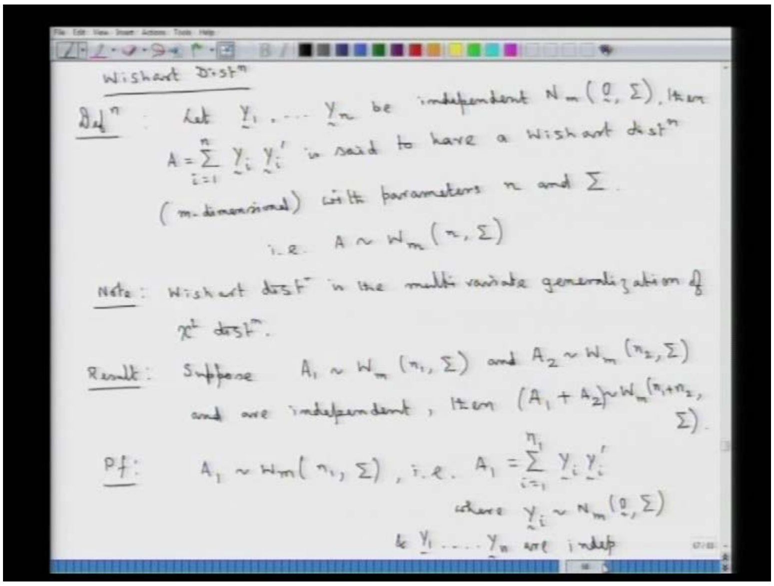
So, said to follow it p-dimensional Wishart distribution with parameters n minus 1 and sigma. We write it in the following way that this n minus one S n minus 1 this follows a Wishart distribution p-dimensional with parameters n minus 1 and sigma. So, this is how Wishart distribution anyway is written. Now, this of course, we are going to prove that this particular quantity n minus 1 S n minus 1 which is the sum of square and cross products matrix the random matrix as we really got it Wishart distribution on the degrees of freedom n minus 1 and. The associated covariance matrix sigma as cleaned in this particular statement we are of course, going to prove it. And we also make a note of this particular fact which we also will be proving that the two statistics furthermore this X bar and S n minus 1 are independently distributed.

So, these are the two important things about the distribution of the sample variance covariance matrix S n minus with a deviser n minus 1 which also happens to be the unbiased estimator. And this X bar and S n minus 1 are going to be independently distributed. Of course, in order to derive this particular result that this really has got a Wishart distribution and this are independent .We would have to talk about what Wishart distribution actually is and its property is (( )) actually move on to that. Let us try to see this result basically is the result which talks about the sample variance covariance matrix form a multivariate normal distribution and then the distribution of the associated statistic X bar and S n minus 1.

Let us try to also tally this results with the type of results that we usually have in case of univariate normal set up. What is result that we have for the univariate normal setup .If we have X 1 X 2 X n a random sample from univariate normal distribution with mean as mu and variance as sigma square, then X bar has got a normal distribution with mean as mu and variance as sigma square by n .What we have here, we have for the multivariate setup a result which is similar to that particular univariate result. Now, the mean there for the univariate normal which was mu was equal to the mean of that associated distribution and the variance of that distribution in case of univariate normal sigma square by n .What we have here in the multivariate setup is sigma matrix divided by n.

Its result which is similar in nature to that particular result and what is a result that we have for the in variate normal a setup, when we talk about estimation of sigma square quantity. Well if we had in such a situation S n minus 1 defined as one upon n minus 1 summation X i minus X bar whole Square. Then the univariate distribution theory tells us that n minus 1 S square by sigma square that follows chi square distribution on n minus one degrees freedom. So, that type of result basically is going to be generalized here in terms of a Wishart distribution which distribution is corresponding to random matrix, because this n minus 1 S n minus 1 which is given by this is the random matrix. We cannot talk about its distribution being chi square and random variable, because that is basically distribution of a univariate random variable. We basically are generalizing that chi square distribution to multivariate set up and we are what we are getting is a Wishart distribution. So, the result is similar nature to that and Wishart distribution in that sense basically the generalization of the chi square distribution in the multivariate setup.

(Refer Slid e Time: 14:06)



So, let us first define what is Wishart distribution how we are going to have Wishart distribution. Let us talk something about a Wishart distribution. Let us first give the definition, let Y 1 Y 2 Y n be independent multivariate normal .Let me not use notation n here, let use in notation N here. So, Y 1 Y 2 Y n suppose is are independent normal m dimension with mean vector as null vector and a covariance matrix as sigma matrix. Then the following quantity A which is summation i equal to 1 to upto n Y i Y i transpose is said to have a Wishart distribution which is going to be off the dimensional of the order of this particular matrix which is n, because we have each of these multivariate normal as m dimensional. So, Y i Y i transposes going to be an m by m random matrix.

So, this is and m dimensional Wishart distribution with parameters m and sigma .We write it as A following a Wishart distribution with parameters this is not m this is with parameter n and sigma. This n is associated with the number of normal random vectors that is what we have Y 1 Y 2 Y n. So, this on m dimension with parameters n and sigma. This sigma thus is associated with the variance covariance matrix of the constituent variate normal distributions, and n is also referred to as the degrees of freedom corresponding to such a Wishart random matrix. Degrees of freedom in the sense that we have Y 1 Y 2 Y n being independent multivariate normal with a mean vector zero and the covariance matrix sigma .Note that each of them has got the same covariance matrix sigma and in such a situation if we look at the sum of this random matrices Y 1 Y 1 transposes random matrix Y 2 Y 2 transpose is another random matrix.

So, Y n Y n transpose is another random matrix all of them are going to be independent. We have with the summation of n such independent random matrices and that is what is going to us this degrees of freedom n here and. we read it in the following way that a follows a Wishart distribution with a on m dimension with n degrees of freedom and with an associated variance covariance matrix as sigma .Now, I say that this Wishart distribution is multivariate generalization Wishart distribution is the multivariate generalization, of a chi square distribution why is that , because suppose you consider the univariate setup in the univariate setup we will say that Y 1 Y 2 Y n are univariate random variables, independent. Each having an univariate normal distribution with mean zero and the variance equal to sigma square.

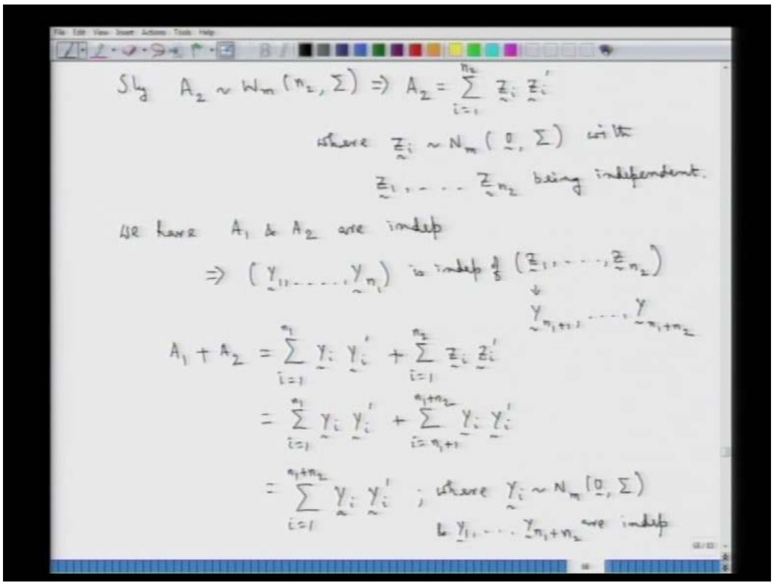
Then if we look at summation Y i square in the case of univariate distribution this Y i Y i transpose both of them are univariate random variable. So, they have one component. So, Y i Yi prime in such a situation will be just Y i square. If we have Y 1 Y 2 Y n uni variate random variables independent normal zero sigma square .Then summation Y i square will have what a sigma square chi square on the degrees of freedom which would be actually the number of independent random variables in that particular summation of whose squares are looking at. So, that in such a situation summation i equal to 1 to n Y i square will have a sigma a square chi square distribution on n degrees of freedom.

So, it is in this way similar to that of a chi square distribution where we are not looking at not Y i squares, but we have multivariate random vector Y 1 Y 2 Y n. And we are looking at Y i Y i transpose and thus this is now having a distribution which is the distribution of random matrix which is a Wishart distribution this m if we had univariate distribution then , this m would at been one this n was degrees of freedom of a chi square distribution. Now, this is the degrees of freedom associated with a Wishart distribution and with the same sigma square of course, we will requiring that and hence this is the associated variance covariance matrix. Now, two simple results concerning a Wishart distribution other following two results. Suppose we have its result concerning the sum of two independent Wishart distributions. Suppose A 1 follows a Wishart distribution m dimension with n 1 degrees of freedom and an associated variance covariance matrix sigma and we have another Wishart distribution with the same dimension m a different degrees freedom n 2 and the same associated variance covariance matrix. Suppose we have A 1 following a Wishart at this A 2 following a Wishart n 2 sigma and are independent .Then, the sum of the two Wishart distribution which is another random matrix. So, this A 1 plus A 2 which is sum of the two Wishart distributions will also follow a Wishart distribution with the same dimensionality .The degrees of freedom n 1 plus n 2 ,and the same associated variance covariance matrix. So, this random matrix A 1 plus A 2 thus also has a Wishart distribution .Once again one would recall such that such a similar result holds for chis distribution which comes from random sampling from univariate normal distribution.

Now, how to prove this particular result, now this result can be proved in various using a distribution for using probability density function. The joint probability density function Wishart distribution using characteristics function of Wishart distribution using characteristics function of Wishart distribution what we will look at, because we are not getting looked at what are what is characteristic function of Wishart distribution etcetera. So, what we will do is to just use the definition of Wishart distribution may order to prove the additive property of this Wishart distribution. Now , this A follows A 1 follows Wishart m n 1 sigma .Now ,from the definition what we can say is that A 1 thus can be written as summation i equal to 1 to upto n 1, because it is on n 1 degrees of freedom. There are n 1 such the random vectors Y i.

This is Y i Y i transpose, where each of this Y I's follow a multivariate normal with dimensionality as the dimensionality of the Wishart and mean vector as null vector and. A covariance matrix as the covariance matrix associated with the Wishart distribution which is n 1 and this Y 1 Y 2 Y n are independent. That is what is the definition of the Wishart distribution.

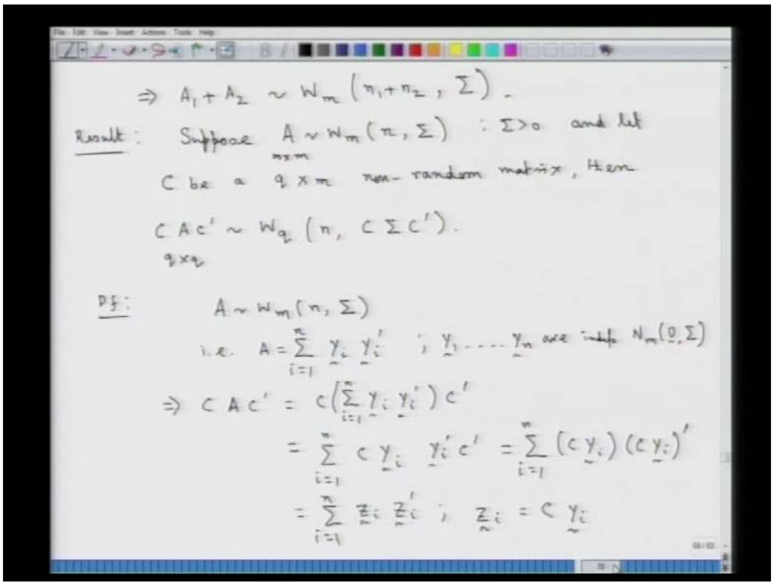
(Refer Slide Time: 22:42)



Similarly if we look at A 2 we have A 2 to follow a Wishart distribution m n 2 sigma. So, this would imply that A 2 is of the form that it is summation i equal to 1 to upto n 2 of Z i Z i transpose where ,each of these Z i s follow a multivariate normal m null and the same covariance matrix and sigma with this Z 1 Z 2 Z n 2 being independent. So, that is the definition of the Wishart distribution once again. So, we will have A 2 to have a this form where the component Z i is they have a multivariate normal with mean vector zero and the covariance matrix sigma and this is Z 1 Z 2 Z n are independently distributed. Now, we are given that this A 1 and A 2 are independent. So, this would imply that this set Y 1 Y 2 Y n is independent of this set which is this is Y 1 Y 2 Y n 1 and this is independent of the other set of random variables which make up this A 2 why is this, because A 1 is given through this Y 1 Y 2 Y n 1 is this random vector and A 2 is given by this set of random vectors Z 1 Z 2 Z n 2.And since A 1 and A 2 are independent this set of multivariate random vectors is independent of the other set of random vectors this Z 1 Z 2 Z n 2. Now, we redefine this quantities say I redefine this as Y n 1 plus 1.So, no problem in redefining that I will call Z 1 to be Y n 1 plus 1 Z 2 to be Y n 1 plus 2 and this Z n 2 to be Y n 1 plus n 2. So, we are just making this Y 1 Y 2 Y n 1 and this Z Z 1 Z 2 Z a Z 1 Z 2 Z n 2 as Y n 1 plus 1 upto Y n 1 plus 2. So, if we under such a situation consider what is this A 1 plus A 2 now A 1 plus A 2 from the definition of this Wishart distributions was i equal to 1 to n 1 Y i Y i transpose and A 2 in terms of Z first is i equal to 1 to n 2 Z i Z i transpose. We have redefined them in terms of Y i. What we can write is this is i equal to 1 to n 1 Y i Y i transpose this plus summation i equal to n 1 plus 1 to n 1 plus n 2 of Y i Y i transpose.

So, this is written in terms of a single summation as i equal to 1 to n 1 plus n 2 Y i Y i transpose where, the characteristics of these element which are there in this summation of order n 1 plus n 2 where, these Y i is follows a multivariate normal m dimension with mean vector as null vector and covariance matrix as sigma and Y 1 Y 2 Y n 1 plus n 2 are independent. We have this particular random matrix summation of n 1 plus n 2 random matrices where, each of the constituent vector Y i is now have got multivariate normal zero sigma and they are all independent. So, from the definition of the Wishart distribution which we had stated in the last slide here. That if this is the set up that they are independent multivariate normal .Then this quantity we will have Wishart distribution with the associated parameters. Thus this quantity here what we have is summation of such quantities which are there in the definition of the Wishart distribution.

(Refer Slide Time: 27:56)



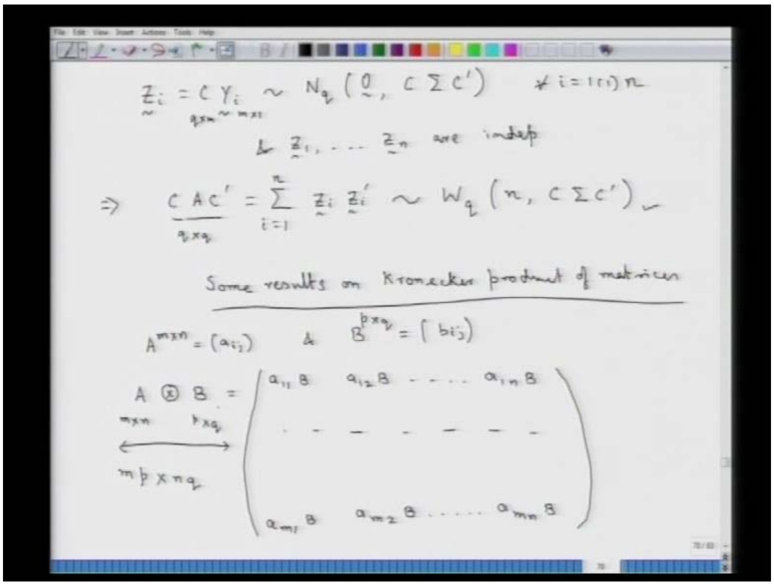
So, this would imply that A 1 plus A 2 will follow Wishart distribution m dimension with parameter as n 1 plus n 2 the diffuse of the freedom of Wishart distribution and the associated variance covariance matrix as sigma. This proves this result. Now, let us look at another simple fundamental result about Wishart distribution. Suppose, we have A to follow a Wishart distribution Wishart m n sigma, sigma is a positive definite will not write it again and again tat it imply that we are not dealing with a singular sigma matrix will always be looking at sigma to be positive definite. Suppose we have a following such a Wishart distribution and let C be a q by m non-random matrix. Once we say that A has got a Wishart distribution m n sigma this is A has this is a random matrix which is of the order m by n.

So, its square matrix and this C be a q by m not random matrix of constants essentially. Then, C A c prime will also have a Wishart distribution. Now, this C A c prime is going to be random matrix, because A is a random matrix of order m by m. So, the order of this C A c prime is q by q. This is q by q random matrix this would have a Wishart distribution q dimension on the same degrees of freedom as the degrees of freedom of the underlying Wishart distribution which is n and an associated variance covariance matrix now as C sigma C prime. This also is a very fundamental result. Once again we will look at proving this particular result without using any further properties of a Wishart distribution like it is a pdf or it is characteristic function what we will simply be using is the definition once again of the Wishart distribution.

So, what do you have we have A following a Wishart distribution Wishart m n sigma. That is from the definition Wishart distribution A is given by the following that it is summation i equal to 1 to n Y i Y i transpose where, Y 1 Y 2 Y n are independent identically distributed multivariate normal m dimension with mean vector as null vector and covariance matrix as sigma matrix . That is the definition. So, if we now look at the quantity whose distribution is desired to be obtained. this would imply that C A c prime is C A is summation i equal to 1 to n Y i Y i transpose times C transpose. So, we pre-multiplied this summation with C and post multiply with c prime. What we have is the following that this is i equal to 1 to n C Y i and this is Y i transpose c transpose.

So, this one can write as summation i equal to 1 to n. Let me just write it as C Y i and this can be return as C Y i transpose. Now, what let us write this as in a new notation Z i Z i transpose where this Z i is a new random vector which is C times this Y i random vector. Now what it is a what is that special about this Z i quantities.

(Refer Slide Time: 32:09)



So, this Z i which is C times Y i .C is a non-random matrix from the properties of a multivariate normal distribution what is the distribution of this Z i which is C times Y i. Now C is q by m matrix this Y m by one vector. So, this is going to have a multivariate normal distribution q dimension with a mean vector the previous mean vector of Y i was null and hence this also remains null and. Then the covariance matrix of C Y is C covariance matrix of Y i times C prime. This is C sigma C prime.

Now, what we had about Y i was that Y 1 Y 2 Y n were independent random vectors and so, will be Z 1 Z 2 Z n ,because Z i vectors are derive from this Y i vectors. What we also have this is for every i equal to 1 to upto n and Z 1 Z 2 Z n are independent. So, what is that we have this C A C prime this would imply that is C S C prime which is nothing, but little in terms of this multivariate normal distributions i equal to 1 to n Z i Z i prime. Now, these Z i quantities are random vectors which are independent normally distributed which this with this as the mean vector and C sigma C prime as it is associated covariance matrix. So, that would imply that this quantity is going to follow Wishart distribution with what dimension with the dimension of this random matrix which is q by q. And, hence this is a Wishart q on what degrees of freedom. Degrees of freedom is the number of independent random vectors here which is n.

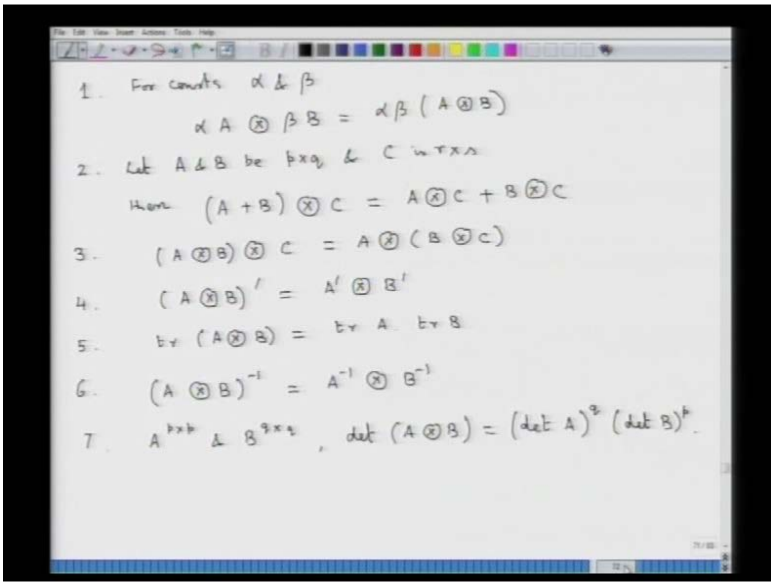
This is the degrees of freedom as n and the associated covariance matrix of the Wishart distribution is the covariance matrix associated with the constituent multivariate normal distribution covariance matrices. So, that is C sigma C prime. So, this proves this previous result which we had stated there if that if A Wishart distribution and q is non-random matrix. Then C A C prime has also what this particular Wishart distribution on this n as the degrees of freedom and C sigma C prime as its associated covariance matrix. Now, before will proceed further we will be requiring some elementary concepts on chronica product. Will just state these results its will not be proving any of this results. I will just state this result, because these chronica product results would be use heavily in further improving for the properties of Wishart distribution. And as also when we try to prove the quantity of interest which we had stated that the distribution of X bar was mu the distribution of n minus 1 S was Wishart and X bar n minus 1 S or X bar and S are independently distributed.

So, in order to prove that we would be requiring these results concerning the chronica product. Let me just introduce what is a chronica product and what are the results that we are going to state here. Some results on chronica product of matrices. Suppose, we have A and m by n matrix comprising of elements a i j and let us have b to B p by q matrix of elements b i j. Then A chronica product this is the sign for the chronica product. I will be using this sign for chronica product hence forth. So, A chronica product B this is m by n matrix this is p by q matrix.

So, this is going to be given by the following matrix, which is a 1 1 which is a scalar quantity that multiply by B , a 1 2 multiplied by the B matrix a 1 n multiplied by this B matrix. And that is all the elements of a are intern multiplied by the same B matrix. This is a m 1 that multiplied by the B matrix, a m 2 multiplied by the B matrix and this is a m n multiplied by this B matrix. So, this is how we define a chronica product between two matrices a and b .Now what is the dimension of this particular matrix if you look at what is happening when we are looking at the chronica product of a with chronica product of a with the chronica product B .Then what is happening is that each of these each of these are matrices which is of the order p by q.

So, we have p by q matrix here one first second then n such p by q matrices augmented one after the other. The dimension of this particular A chronica product B matrix which is given on the right hand side is going to be m p times m q. This is what we are going to have we look at the chronica product of A and B. Given this A B two matrices this is how the chronica product is defined.

(Refer Slide Time: 38:15)

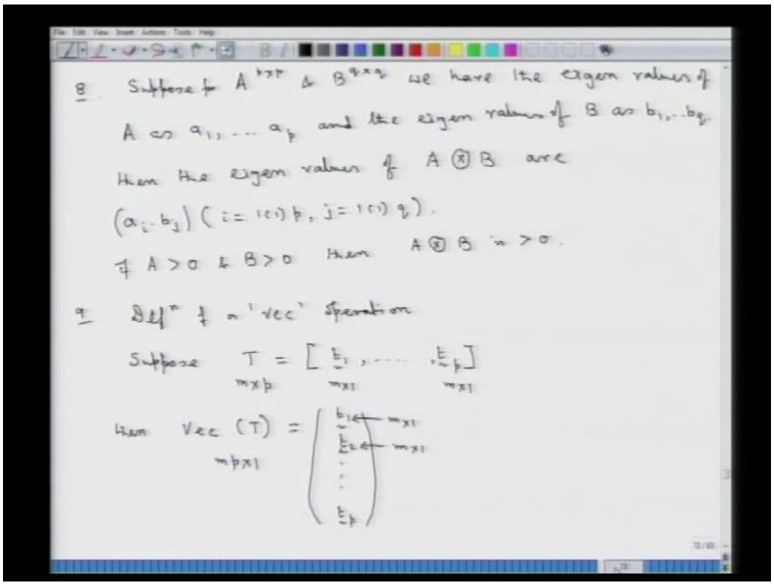


Now, I am going to state these results .The first that I will write is for constants for scalar constants alpha and beta say and A and B two matrices as we had defined. This alpha A chronica product beta B is going to be given by alpha beta times A chronica product B. So, if we have two constants scalar constants alpha and beta alpha B multiplied with A and then it chronica product is taken with matrix beta multi B multiplied by the beta scalar constant then that is this quantity. Let A and B be two matrices p by q and let me have C as r by s. Then, this A plus B this is the just sum of the two matrices chronica product this matrix C is nothing, but A chronica product C plus B chronica product C. Now, the third result that we might be requiring is that suppose we have A B C three matrices such that the following multiplications are possible.

So, if we have A chronica product B matrix being multiplied in the ion the terms of chronica product with another matrix C then, what we can say that this is equivalent to A chronica product of B chronica product C. Now, the fourth results is what if we have A chronica product B the matrixes transpose then that would be given by A transpose chronica product B transpose. We have also the following result for the trace of two chronica product the chronica product of two matrices. If we look at trace of A chronica product B matrix this all this results are elementary and simple to prove also. We that will be trace of A times this is scalar quantity times the trace of B matrix. The result concerning the inverse of chronica product of two matrices is the following that suppose we have A chronica product B inverse of that being taken.

Now, when we talk about this inverse this is going to be A inverse chronica product B inverse. Here of course, A is not p by q or B is not m by n . Both this A and B are square matrices square matrices which are non-singular and hence we will assume that in such a situation when we are looking at A chronica product B is inverse to be return as A inverse chronica product B inverse. Then both this A and B matrices are non-singular. We will also be requiring this result here, that suppose A is p by p and B is say q by q then, we will have the determinant of A chronica product B to be given by. So, this is a p by p and this is q by q. This would be given by determinant of A raise to the power order of the B matrix this multiplied by the determinant of B raise to the power of order of the A matrix .So, this is this p. Now, remove on to stating three more results.

(Refer Slide Time: 42:21)



This number eight would be suppose this A p by p suppose for p by p A and B say q by q we have the eigen values of A as a 1 a 2 a p and the eigen values of B as b 1 b 2 b q. This a 1 a 2 a p are the eigen values of the A matrix and b 1 b 2 b q are the eigen values of this B matrix. Then, the eigen values of A chronica product B matrix. Now, what is the order of this A chronica product B matrix the order of this A chronica product B matrix would be p q rows and p q columns. We will have p q eigen values of this A chronica product B matrix and these are the basically given by then the eigen values of these are a i times b j i equal to 1 to upto p and j equal to 1 to upto q. This a i b j is are a i times b js.

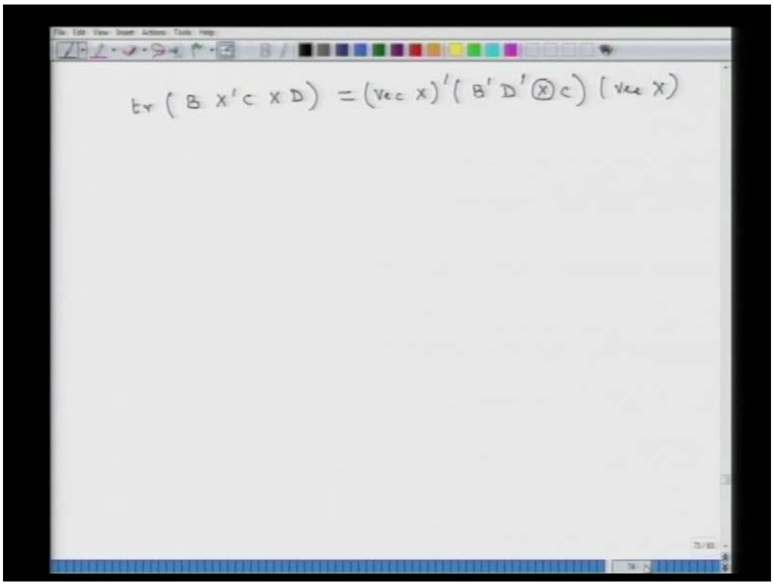
So, these are going to be the eigen values that are associated with A chronica product B matrix there are p q of those eigen values here .Now, note that if we have A to be positive definite B to be positive definite. So, will be A chronica product B matrix, because each of these eigen values if A is positive definite are greater than zero and if B is positive definite we will have all these to be also positive definite to be greater than zero b 1 b 2 b q and .Hence these product a i times b j is which are associated with the eigen values of a chronica product B matrix. Those are also going to be greater than zero strictly. Now, if you have one to be positive definite one to be positive semi-definite what is going to happen A chronica product B .If A is positive definite all these a i is a greater than zero A is positive semi-definite.

This b j is are greater than or equal to zero. So, we can at the most say that a i times b j are all greater than or equal to zero. And hence the A chronica product B matrix would in such situation be positive semi-definite. So, that is what if A is positive definite and B is positive definite then A chronica product B is also positive definite. Now, the last result that we will be stating for this chronica product of matrices is the following, but before that we need to define a vec operation what is that. So, definition of a vec operation first we will define this and then in terms of this vec operation we will have the last result in this section being stated.

Suppose we have we have A matrix T say I take that to be a of order n by p which is comprising of these column vectors t 1 t 2 this is p. This is t p. Each of these are m by one column vectors. We have this matrix t which has m rows corresponding to the dimension of each of this t i vectors .Then, this vec of this T matrix is given by the following that we stack up all these columns one below the other. What we have done in this vec operation is a following that we had that this T matrix to be comprising of p m dimensional column vectors.

So, what we have done in the vec is the first m components. This is and m by 1 component and so will be this t 1 this is the second stacking here which is also m by 1. Then the third t 3 would be coming. So, we place t 1 at the top then t 2 t 3 and like that. We make a column vector from this particular n by p matrix and this thus would be a column vector of the order m p by 1.We will have all these being stacked one over the other and thus giving us this vec of this particular T matrix. Now this is particular this would be particularly useful when we are immediately going to define what is a matrix normal distribution will use this (()). Now if this is vec operation then the following result concerning the chronica products also hold true.

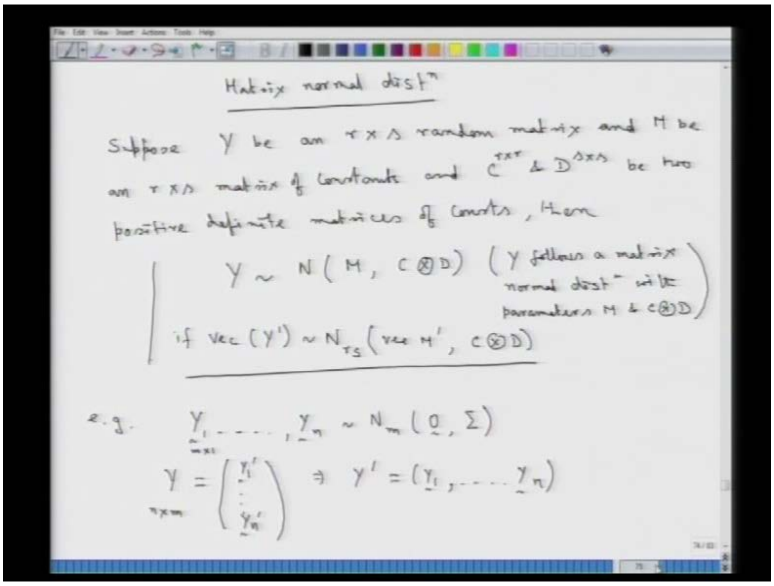
(Refer Slide Time: 47:58)



Suppose we have let me not put any restriction on these matrices. We will have trace of the following quantity where this is B X transpose C X D. Then, trace of this quantities going to be given by vec of X transpose times B transpose multiplied by this D transpose chronica product this C matrix that multiplied by this vec of X. So, this is what is going to happen. Now, this B X transpose C X also and D are matrices such that this multiplication, because this is just matrices multiplication operation they are defined. As long as they are defined for any such B X C and D matrices this result (( )) hold true that such the trace of this can be written in terms of this. Now, with this particular introduction about chronica products.

Let us look at the definition of matrix normal distribution and how we can define a Wishart distribution that .We have define through multivariate normal random vectors that also can be defined through a matrix normal distribution and then of course, I will show that the two definition of we start distribution are equivalent. Why this is required this is basically required in order to prove there is a result which we have stated sometime back that X bar and S are independent and the distribution of S is Wishart distribution n minus 1 S is Wishart we will be using the definition of a matrix normal distribution in such a situation.

(Refer Slide Time: 49:49)



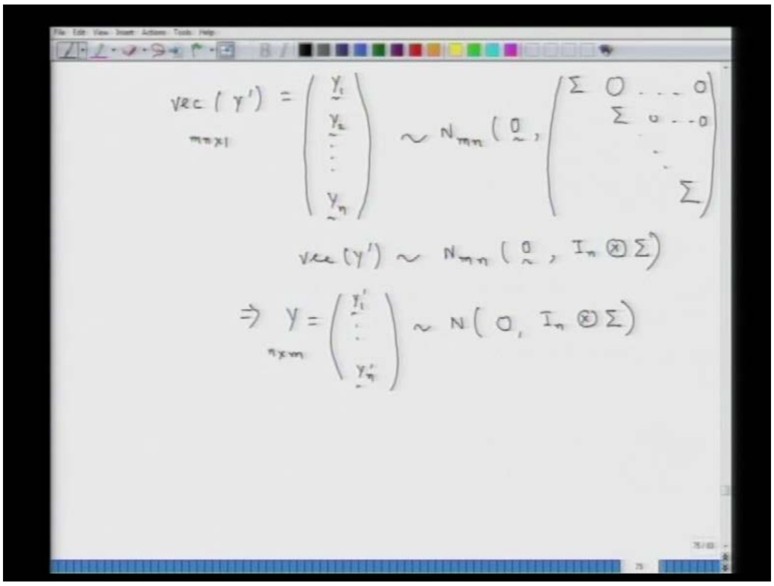
Let us first define a matrix normal distribution .Suppose we have Y be an r by s random matrix and M be an r by s matrix of constants and .C and D let us make C to be r by r and D to be s by s be two positive definite matrix be two positive definite matrices. Then they of course, are positive definite matrices of constants. So, we have 1 2 3 4 such matrices. Now Y is an r by s random matrix m is r by s matrix of constants C and D are also two non-random matrices of constants this is r by r and D is s by s. These two matrices C and D are positive definite matrices m may or may not we are not actually concerned about any special property about this m matrix is a rectangular matrix it is importantantly of that is same dimension as that of this particular matrix.

Under such a setup this random matrix is C to have a matrix normal distribution I just write it as N, but this should be red as a matrix normal. Because this Y is a matrix of matrix r by s matrix its random matrix. And hence, the we have this being defined as n m this is that matrix and C chronica product D. If I will just write it in bracket how we read it Y follows a matrix normal distribution with parameters M and C chronica product D. We will know what these parameters actually represent for the case of this matrix of normal distribution once we give the definition of this matrix normal distribution .Then Y will follow a matrix normal distribution this if vec of Y prime follows a multivariate normal of dimension r times s with a mean vector as vec of M prime and a covariance matrix which is C chronica product D. This what gives us the definition of a matrix normal distribution Y is said to have a matrix normal distribution with a matrix m as one set of parameters and C chronica product D as the other set of parameters.

If we have vec of Y prime now what is the dimension of vec of Y prime Y is r by s. We can think that Y is comprising of s column vectors each of dimensions r. So, if we make vec of Y prime from their then it is going to lead us to and r by s random vector. And that is what is going to have a multivariate normal distribution r s dimension with vec of m prime which is going to have the same dimension as vec of Y prime which is r s as its mean vector and C chronica product D as its associated covariance matrix .Now for example, suppose we have Y 1 Y 2 Y n each of them following a multivariate normal m dimension with a mean vector zero and covariance matrix as sigma matrix. Then, let us define Y matrix to be the following Y 1 prime Y 2 prime Y m prime.

So, this is going to be a matrix now each of these are m by one vectors. This is a row vector of how many columns m. This is going to be an n by m matrix. Now, from here if we look at what is Y prime Y prime is going to be Y 1 Y 2 Y n. So, if we look at now vec of Y prime and let us see what is the distribution of that .Now Y note that this Y is a random matrix because it is comprising of these rows each of these are random variables random vectors rather.

(Refer Slide Time: 55:33)



Now, vec of Y prime is thus going to be a column vector how is that going to be formed. The first set of entries is going to be this Y 1 which is m dimensional, the second set of entries would be this Y 2. So, we are stacking one vector after the other and this is going to be this Y n. So, this vec of Y prime is going to be m n cross one dimensional random vector. Now what is the distribution of this you we have taken actually this each of these Y is to be this and they are independent say. Suppose I take Y 1 Y 2 Y n multivariate normal m dimensional with a mean vector as zero and covariance matrix and sigma and they are independent. Then, what is going to be the joint distribution of this is just stacking of this multivariate normal distribution and hence this is going to multivariate normal m n dimension what would be the mean vector each of the mean vectors are null vector.

So, this is going to be null vector of dimension m n and what is going to be the covariance matrix of this vec of Y prime. Now note that each of these Y i is Y 1 Y 2 Y n has covariance matrix sigma. This is going to be a matrix which is going to have sigma matrix has its block diagonal matrices and what are the half diagonal matrices. Now this of half diagonal matrix the one-twelfth position block diagonal matrix would be the covariance matrix between Y 1 and Y 2. Now Y 1 and Y 2 are independent multivariate normal random vectors and hence the covariance will be these null matrices.

So, these are all null matrices here and we can write that compactly as m n as I n chronica product sigma I n is an identity matrix of order n. We have this vec of Y prime to follow this. So, we had in the definition of a multivariate matrix normal distribution that we will say that this random matrix Y has got a matrix normal distribution with these set of parameters if vec of Y prime as got this result distribution a multivariate normal distribution and hence that is what we have here. So, this would imply that Y the random matrix what we had defined as Y 1 prime Y 2 prime Y n prime. This was that n by m random matrix this is going to have a matrix normal distribution with a null matrix m here this is taking the place of N there and I n chronica product sigma as the second set of parameters.

We see that how we actually get a matrix normal distribution from a random sampling through rather a random sampling from multivariate normal distribution. If we have Y 1 Y 2 Y n all multivariate normal zero sigma. They are also independent then, under such a situation if we frame such a matrix Y then this matrix Y will have a matrix normal distribution with the associated m matrix their as a null matrix. And C as I n and D as this sigma the variance covariance matrix of the associated normal distribution. Next time we are going to use this particular result and this definition of matrix normal distribution in order to given a alternate definition of Wishart distribution and derive results of importance. Thank you.