Applied Multivariate Analysis

Prof. Amit Mitra

Prof. Shramishtha Mitra

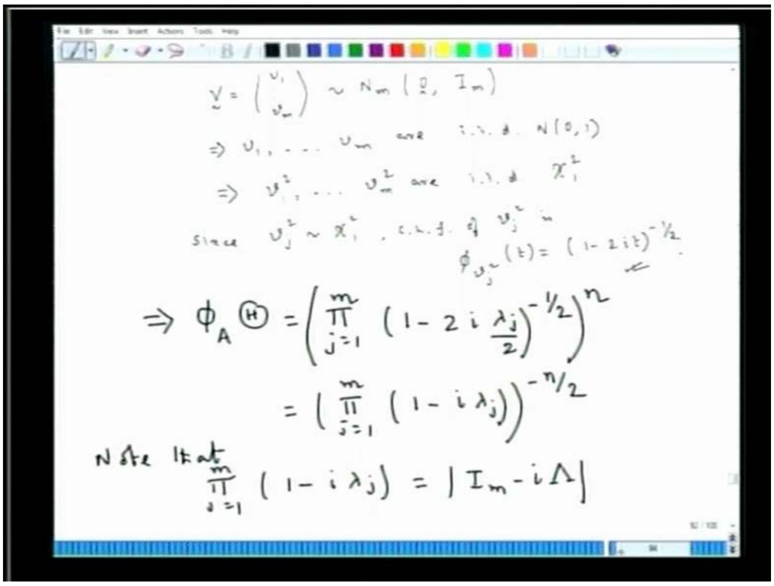
Department of Mathematics and Statistics

Indian Institute of Technology, Kanpur

Lecture No. # 12

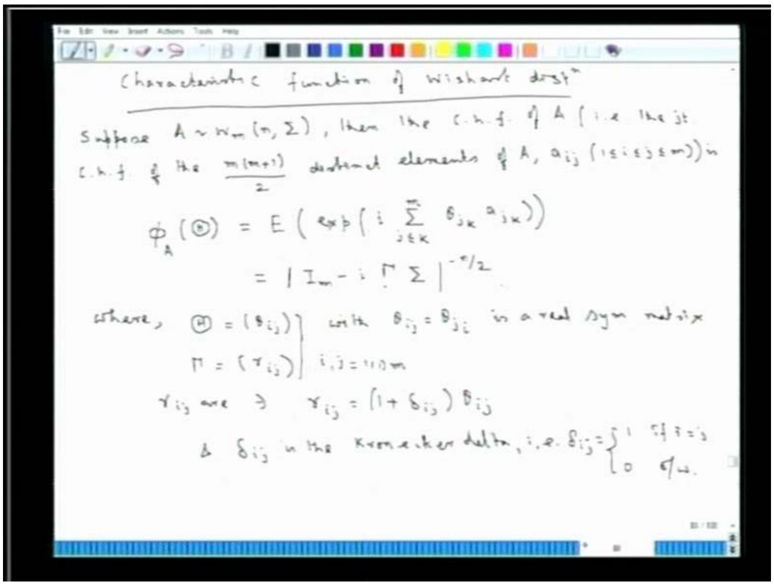
Wishart Distribution and it's properties – II

(Refer Slide Time: 00:20)



In the last lecture, we had started actually proving the characteristic function or deriving the characteristic function of a Wishart distribution. Let see, where we were actually in the last lecture.

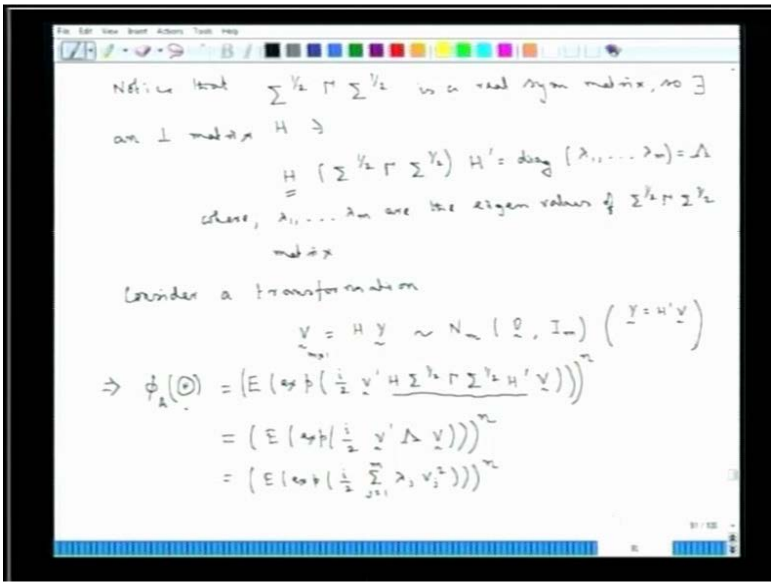
(Refer Slide Time: 00:32)



So, we were looking at proving this particular result that, suppose A has got a Wishart distribution, m-dimensional with degrees of freedom as small n, and associated variance, covariance matrix as sigma. Then the characteristic function of A the random matrix, it is a symmetric matrix.

So, we are looking at the joint characteristic function of m into m plus 1 by 2 distinct elements of A given by denoted by a i j. So, this quantity is what is giving us the joint characteristic function or the characteristic function of the Wishart distribution, which we were trying to prove that it is determinant of I m minus i square root of minus 1 times gamma into sigma determinant of that, whole raise to the power minus n by 2; where gamma matrix was given in this form, and sigma matrix was the associated variance, covariance matrix.

(Refer Slide Time: 01:30)



So, in proving in this particular result we had come up to this point that, we had shown that this characteristic function is of this form, that it is expectation of E to the power i by 2 then, summation j equal to 1 to m up to the dimension of the underline multivariate normal distributions of lambda j, v j square, and whole these thing raise to the power n; where this lambda j is; where the eigen values associated with sigma, half gamma, sigma half matrix. So, that this lambda, capital lambda matrix is the diagonal matrix containing the eigen values lambda 1, lambda 2, lambda m, and this v j are independent chi square random variates, because v j squares are independent chi square random variates. Each of these elements v j, which are there in this vector v has got i i d normal with mean 0, and variance equal to 1.

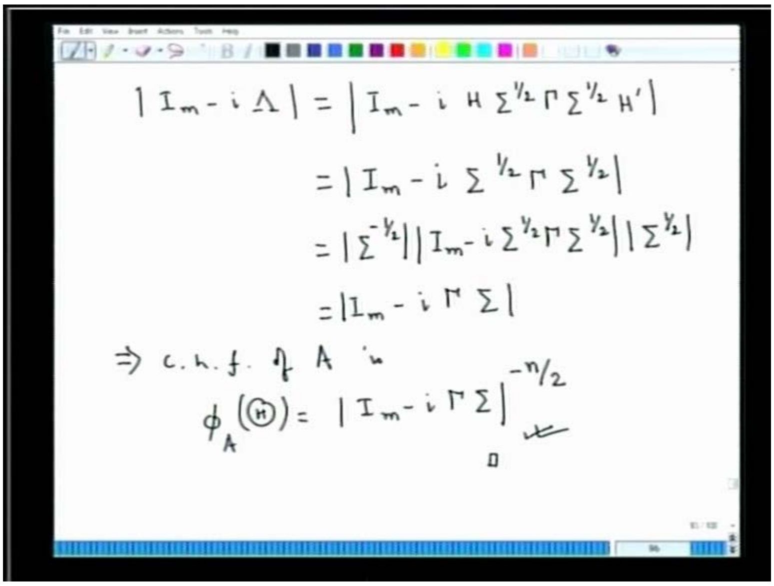
So, they were standard normal variates, and we realized that this v is v 1, v 2, v m; which has multivariate normal I m i with the null vector as it is mean vector, and I m as it is variance, covariance matrix. So, that would imply these things straight away, and this v 1 square, v 2 square, v m square, are i i d chi square 1 random variate, and hence using the characteristic function of a chi square random variate on one degrees of freedom; what we had was this each of the v j square random variates had a characteristic function this.

So, this would imply now that the characteristic function of the Wishart distribution, in which we were interested in which we had expressed in terms of that expectation would now take the form; that this is i equal to 1 to up to m 1 minus 2 i t means this t here, when we look at the form of the expression, that is what we had it is basically, lambda j by 2 was serving the purpose of that t in the characteristic function expression. So, then this will be a lambda let me write this index as j because we have already used this i for the complex number and hence this is this lambda j pi 2, whole raise to the power minus half and then this entire expression is raise to the power n

So, that inside this bracket what we have is this quantity computed for each of these chi square random variates, and that raise to the power n. So, what is this this can be written as j equal to 1 to m 1 minus i times lambda j. These i take minus n by 2 outside, and keep it here, now what is this quantity note that. If we look at this product j equal to 1 to m 1 minus i times lambda j. This can be written in terms of determinant of two diagonal matrixes worth are those this is i m minus i times this capital lambda.

So, that this resultant matrix here will also be a diagonal matrix with elements as 1 minus i times lambda j, and then the determinant of that diagonal matrix would just be the product of the diagonal entries. So, this is equal to this we can make further simplification to this expression, and this determinant of i m minus i times capital lambda.

(Refer Slide Time: 05:25)



Now, remember that this lambda is that diagonal matrix containing the eigen values lambda 1, lambda 2 , lambda n, that is from this expression what we have is this lambda matrix is H sigma to the power half gamma sigma to the power half H transpose. So, we can simply take that and write it here that this is equal to i m minus i times H sigma to the power half then this gamma matrix sigma to the power half into H transpose.

Now, pre and post multiplying by pre multiplying by determinant of H transpose and post multiplying this expression by determinant of H. We can take H, H transpose inside, but here we will have that H transpose H, which will once again be an identity matrix, and here if we multiply by H transpose. Then this becomes an identity matrix, and when if we post multiply this by H. So, H transpose H also will become an identity matrix.

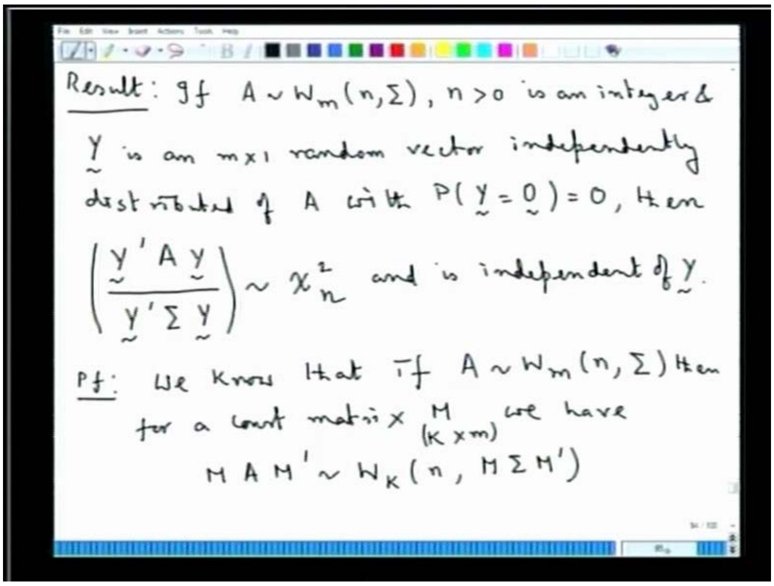
So, what we will be having is i times sigma to the power half gamma, sigma to the power half; now this expression, let me write this expression; this is suppose we pre multiply this now by sigma to the power minus half determinant. This multiplied by I m minus i times sigma to the power half gamma, sigma to the power half, and then post multiply this by sigma to the power half determinant.

So, we can take this sigma to the power half inside this expression, and sigma to the power minus half from the left hand side, and sigma to the power plus half from the right hand side. So, what we will be having is determinant of I m minus. So, this will be an identity matrix this will be i times gamma matrix that multiplied by this sigma matrix.

So, if we have this expression, which is now from here in the final expression of the characteristic function; which we have this term, and we have shown that this term is determinant of this I m minus i times gamma, sigma. So, this would imply that, the characteristic function of A is finally, given by the form that was desired; that this is at the point script theta matrix. So, this is determinant of I m minus i times gamma into sigma determinant of that that raise to the power minus n by 2. I fill it back as the statement of this particular result this is what we were supposed to prove.

So, phi a script theta was determinant of I m minus i times gamma sigma whole raise to the power minus n by 2, and this precisely we have derived in this particular form. So, that is the desired form of this particular characteristic function of the Wishart distribution. So, the characteristic function of the Wishart distribution can actually be used to prove many of the results, and it is a fundamental concept, and hence we had looked at the derivation in detail of the characteristic function of the Wishart distribution.

(Refer Slide Time: 09:20)



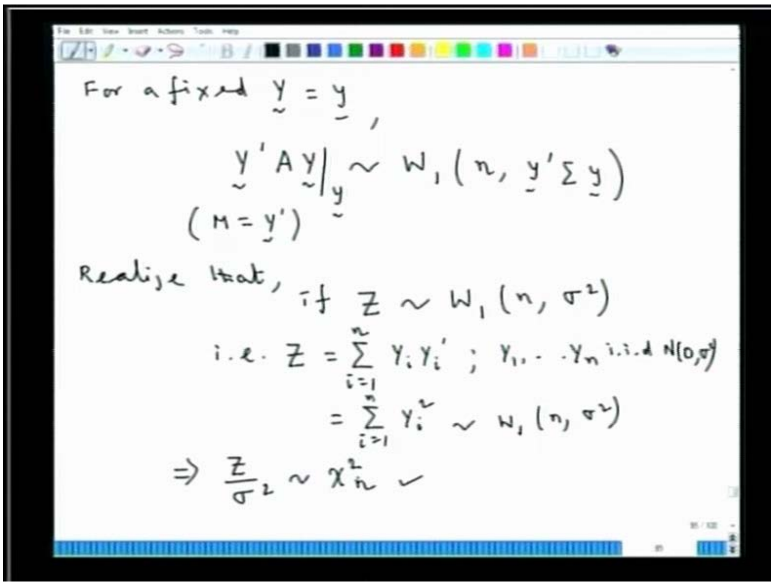
Now, let me go through some important results in the Wishart distribution theory, before we actually move on to defining what is hotelling T squared distribution, and how hotelling T squared distribution is obtained from the Wishart distribution, and a multivariate normal distribution. Let us, first look at this important result; if we have a following a Wishart distribution, Wishart m n sigma with n greater than zero; a positive integer n is an integer, it is the degrees of freedom it is naturally the an integer, because in the definition of a Wishart distribution. We have this n as the number that is associated with the multivariate normal distributions which is associated with that Wishart distribution.

So, if we express A using the first fundamental definition of the Wishart distribution. We can write A as summation Y i, Y i transpose, and that summation is from i equal to 1 to up to n. So, the n, the degrees of freedom is associated with the number of independently an identically distributed multivariate normal distribution; each having a mean, each having multivariate normal m dimensional with mean vector as a null vector, and the covariance matrix as sigma matrix, and hence this of course, is an integer, but for completion, I write as an integer, and Y is an m by 1 random vector independent independently distributed of this random matrix A with this that probability; that y vector takes the null vector is equal to zero. It would be obvious why we take such a condition; then the distribution of Y transpose A Y is going to have a central chi square will that result we had noted last time actually. So, we will proceed with that result.

We will have this written as y transpose A Y this divided by y transpose, A Y transpose, sigma Y. This will follow a central chi square on n degrees of freedom, and is independent of this random vector y. Now note that first, I have written this y transpose A Y the distribution of y transpose A Y is obvious actually, because we had proved a result on Wishart distribution; which said that if A have Wishart distribution; then for a constant matrix A, we will have a sigma A prime to have a chi square distribution to have a Wishart distribution, and if we have certain condition; then that would follow a chi square distribution.

Now, this result tells us that, if we are looking at this particular ratio that Y transpose A Y by Y transpose, sigma Y; that would follow central chi square, and that would be independent of this Y. Now, what we note? You start with is the result that, I was referring to we know that, if we have A to follow a wishart m, n sigma; then for a constant matrix M; which is say k by m order we have M A M prime to follow a Wishart distribution k dimensional on n degrees of freedom, and with the associated variance, covariance matrix as M sigma, M prime we are explicitly proved this particular result that such a result holds true.

(Refer Slide Time: 14:05)



Now, what we will be doing in order to prove the given result is we will assume first that; for a fixed Y equal to this small y. So, this is for a fixed value of this Y equal to y what is going to be the distribution of this Y transpose A Y now since, we are taking y as fixed. It is as, if that Y is given to be small y.

So, at that particular fixed point we will have this distribution the distribution would follow from what we have here in the previous slide; that if we take now M equal to Y then or rather M equal to Y transpose. So, this is basically, in the previous result we are taking M equal to Y transpose. So, what we will be having this as Wishart distribution which had k. Now, what is the order of k; here, k is equal to one, because we are taking Y transpose Y is in vector which is m dimensional.

So, this is going to be wishart with degrees of freedom as n, and what is going to be the variance, covariance term. Here, that is going to be this for a given small y, I should write. So, let me write that as y transpose sigma y. So, this expression for given Y will follow a Wishart distribution one, and y transpose sigma y, because what we have done in the previous result is to just use or rather take m to be equal to y transpose at this particular given small y value.

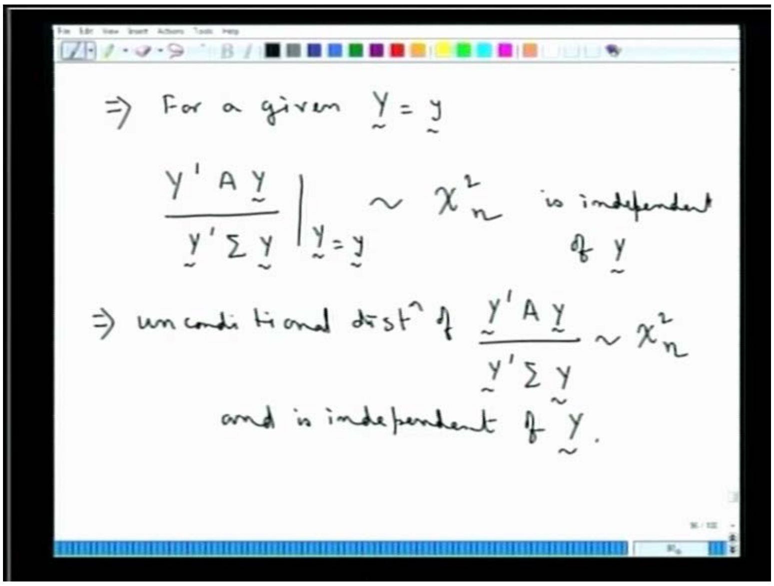
Now, realize the following what is. So, special about a Wishart distribution on one degrees of freedom let me write that as realizing that realize that if we have Z following a Wishart distribution on 1 n sigma square; now since, this is one wishart on one dimension; this is going to be a scalar quantity as you can see here this is y transpose sigma y. So, this is 1 by m this is m by n and this is m by 1 and hence this y transpose sigma y that is actually a scalar quantity.

So, if we have Z following Wishart 1 n sigma square this would imply that Z from the definition of the wishart distribution would be summation i equal to 1 to up to n y i; some other random variable not to be confused with these Y i is here. So, from the definition this is Y i Y i transpose; now what is that about these y i this Y 1, Y 2, Y n. These are scalar random variables, these are going to be i i d normal one dimensional, because the associated Wishart distribution is one dimensional this with the mean zero, and the variance equal to sigma square. That is now, it is equivalent one can actually remove this transpose, because these are scalar random variables. So, this is nothing, but summation i equal to 1 to n Y i square right. So, this follows Wishart 1 n times sigma square, but independently this looking at this particular summation summation i equal to 1 to n, summation Y i square; each of these Y i are normal zero sigma square, and they are independent.

So, this would imply that this Z which is summation Y i square that divided by sigma square will follow what that will follow a chi square central on n degrees of freedom because what we are doing is Z divided by sigma square is nothing, but summation Y i square minus sigma square. So, each of these terms some Y i square by sigma square they have a chi square random variate on one degrees of freedom

So, since they have got chi square on one degrees of freedom, we have the summation of n such independent chi square random variates to have a chi square on n degrees of freedom. So, we will use this particular thing that, if Z follows a Wishart 1 n n, and scalar sigma square; then this Z by sigma square has got a chi square distribution. So, what happens if we use that in this result?

(Refer Slide Time: 19:04)



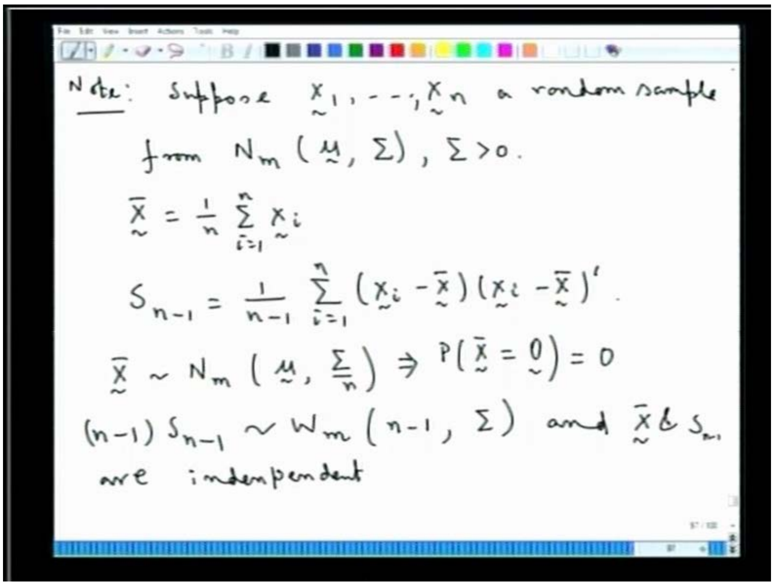
So, this will imply that for a given Y equal to y; this Y prime A Y, this divided by the corresponding sigma square there which is Y transpose sigma Y, I will just put it like given this Y equal to y. This will follow straight forward a chi square of n degrees of freedom.

Now, if we have the distribution of this given Y equal to y; a chi square random variate we note that this distribution that, we get of this quantity given Y equal to small y is independent of this Y. So, whatever be the fixed value of this Y vector at small y; whatever be the fixing vector here small y, this is always going to have the same identical distribution which is a chi square distribution or n degrees of freedom.

So, this would imply now this is basically, the conditional distribution of this given Y equal to y. That is following a chi square n random variate, and that is independent of the fixing small y of this random vector capital y. So, this would imply that the unconditional distribution unconditional distribution of this Y transpose A Y this divided by Y transpose sigma Y. This also is going to have the same conditional distribution; this is the conditional distribution of Y transpose A Y by Y transpose sigma Y given Y equal to small y; that has got a chi square distribution, and that does not depend on the particular fixing vector Y, and hence the unconditional distribution of Y transpose A Y by Y transpose sigma Y is also chi square random variate.

We see that the conditional distribution is the same as that of the unconditional distribution, and hence this basically is independent of this random vector Y and is independent of this random vector y. So, thus proving this important result that, this has got a chi square central of n degrees of freedom, and this random variable here is independent of this vector Y.

(Refer Slide Time: 21:57)

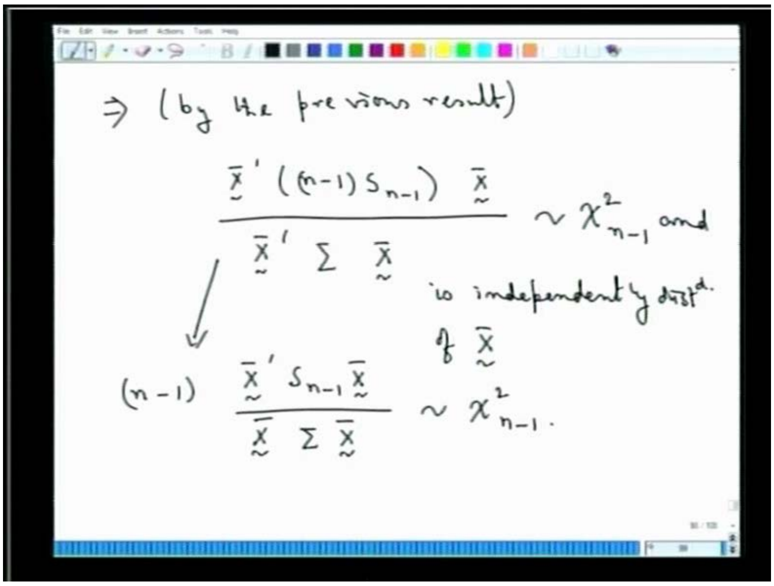


Let see, how way we can use this particular result for random sampling; in case of a normal distribution suppose, we have x 1, x 2, x n, a random sample a random sample from a multivariate normal m dimension the mean vector mu, and a covariance matrix sigma both mu, and sigma are unknown sigma is positive definite. Then we have the two quantities of interest: the two statistics, x bar which is 1 upon n summation of these x i quantities i equal to 1 to n. This is the sample mean random vector, and S say n minus one, 1 upon n minus one, summation i equal to 1 to n x i minus x bar into x i minus x bar transpose.

So, we have these two quantities of interest that x bar, and S n minus one. We have proved in the last lecture, and the lecture prior to that the important result; that this x bar follows a multivariate normal m with the mean vector mu, and the covariance matrix sigma by n, and n minus one, S n minus 1 which is this expression, the sum of squares, and cross product matrix this has a Wishart distribution this has a wishart distribution m on n minus 1 degrees of freedom, and an associated variance, covariance matrix of sigma, and further more we have importantly proved that x bar, and S or S n minus; here one can also write this as S n 1 are independent. So, whatever be it they are going to be independent. So, this x bar, and S n minus 1 are independent. right

So, we will use these two: in order to give to an expression which is very much of interest; we will basically, be using the previous result. Here, now note that if we have got x bar to follow this multivariate normal distribution. This would imply that probability of x bar vector to be equal to null vector; what is that probability that probability is equal to zero, because it is a multivariate normal distribution. So, that random vector taking this vector null vector is obviously zero, and this x bar, and S n minus S n minus one; they are independently distributed.

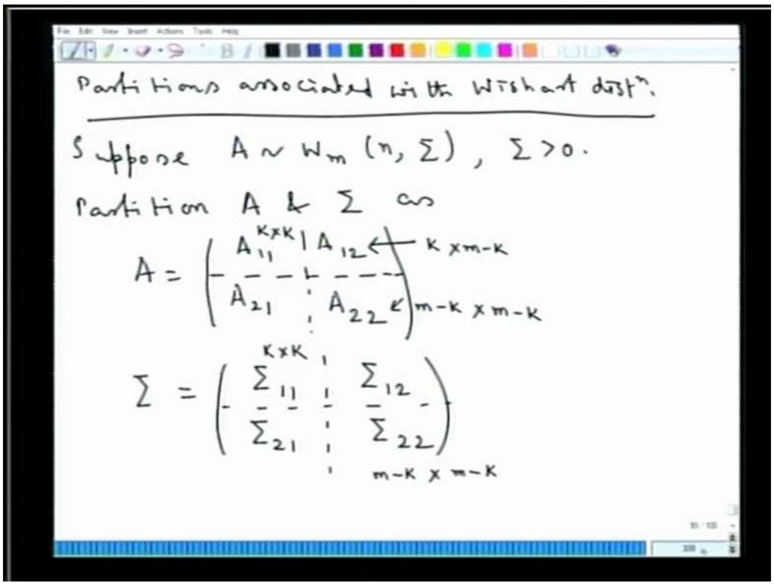
(Refer Slide Time: 25:09)



So, what can we do? We can write that this would imply by the previous result by the previous result that x bar transpose. Now, this has got Wishart distribution; there n minus one, S n minus 1 inverse x bar that divided by x bar transpose, and then what is required here is the sigma inverse. So, this sigma inverse which is associated with the Wishart distribution remains, as sigma inverse this into x bar vector. If we are interested in knowing it is distribution; then this has got a chi squared central on what would be the degrees of freedom the degrees of freedom would be associated with the degrees of freedom of this n minus 1, S n minus 1 Wishart quantity. So, that is a chi square on n minus 1 degrees of freedom, and is independently distributed of the random vector that is x. right

So, this quantity this one can simplify this, and write this as n minus one. So, that this would be there is no inverse there. So, by the previous result this of course, the next result is going to involve the inverses. So, this is going to be this x bar transpose n minus 1, S n minus one. That is going to have a chi square distribution, because let me just go back one slide. So, that this is what we have is Y transpose; the Wishart matrix here Y. So, that would be having this chi square distribution and hence this distribution here this x bar transpose S n minus 1 into x bar. This divided by there is no inverse; here, the inverse will come in the next result. So, this x bar transpose S n minus one, n minus 1 into S n minus 1 x bar; that divided by x bar transpose sigma times x bar. So, that this sigma times x bar, this will follow a central chi square on n minus 1 degrees of freedom.

(Refer Slide Time: 27:56)



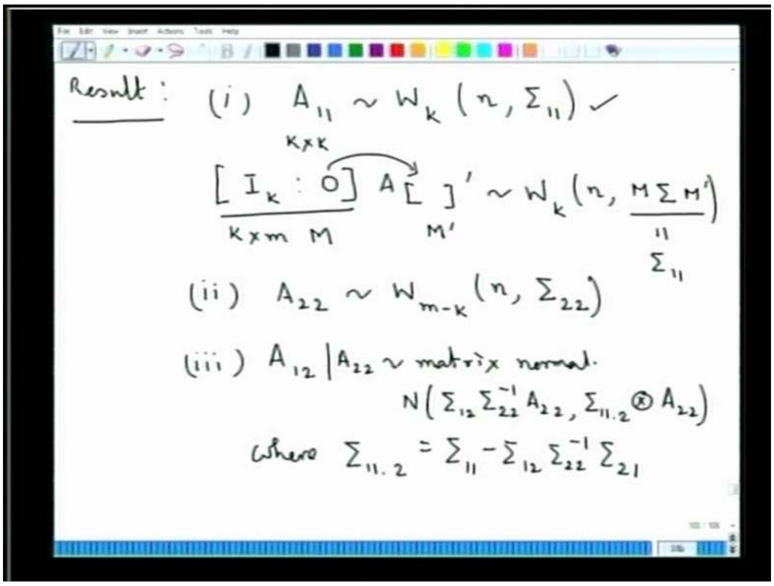
Now, the next step we are going to look at this wishart partitioning partitions associated with Wishart distribution. Now let us, make the following partition in the Wishart distribution which we write as following: suppose, we have A to follow a Wishart distribution m, n sigma sigma of course, associated to be positive definite.

Now, partition A, and sigma as follows partition A, and sigma as suppose we have A to be partitioned as A 1 1, A 1 2. Now, this A 1 1 let us make that A k by k matrix, and this is A 2 2; this is A 2 1 matrix. Now, this has now got the order that it is m minus k by m minus k. Now, the order of this A 1 2, and A 21, are accordingly obtained for example: A 1 2 has got k rows, and it has got m minus k columns, and similarly this has got m minus k rows, and k columns. right

So, this is a partitioning of A; that we have obtained, and rather we have made, we have a similar partition in the sigma matrix; which is sigma 1 1; this is sigma 1 2; let me, put this partitioning like this. So, that we have this as sigma 2 1 that into, and that sigma 2 2 component. So, the partitioning of sigma is as in the partitioning of A, that is this is now a k by k matrix, and this sigma 2 2 is m by k into m by k matrix.

Then, we may be interested in knowing what is the distribution of this sub partition; here, A 1 1 or A 1 1 or some derived form, which involves all the partitioning element of the wishart matrix.

(Refer Slide Time: 30:40)



Let me, write the results for this setup, now the first result: that is what we will be having is A 1 1 will also be having a Wishart distribution k dimensional; now remember that the type of partitioning, that we had made. A 1 1 is the k by k random matrix derived from the wishart matrix. So, this is a Wishart distribution of k on degrees of freedom as the degrees of freedom of the original Wishart distribution, and the associated variance, covariance matrix as sigma 1 1.

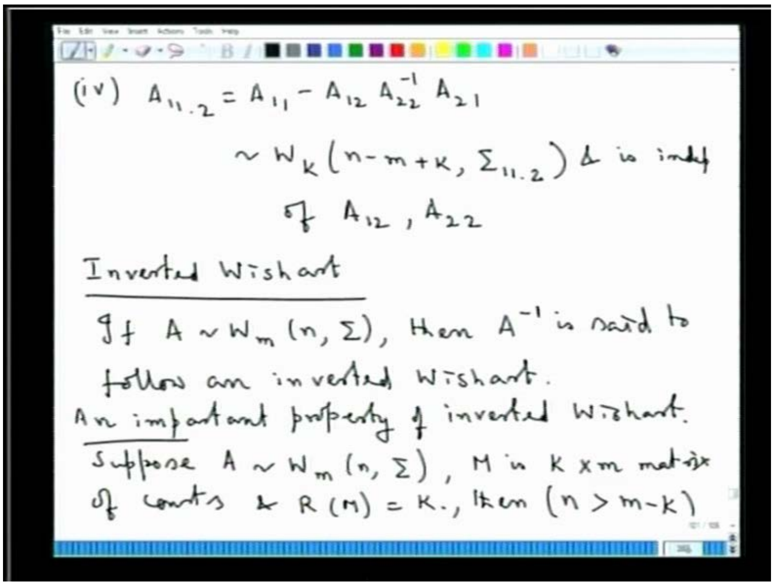
Now, how do we prove that? It is easy actually, all these results can be proved; now say suppose I take an, I k matrix, augmented with a null matrix. So, as to have this as k by m matrix here that A into the transpose of this matrix. So, treating this as the matrix of constant, which is M A M transpose. So, this will follow a Wishart distribution on k dimensions degrees of freedom remaining n, and the associated variance, covariance matrix as M sigma M prime. So, this would be M sigma M prime.

Now, what is this going to be equal to for a sigma; the type of partitioning sigma 1 1, sigma 1 2, sigma 2 1, sigma 2 2; what we have this is going to just be equal to sigma 1 1. So, that this partition A 1 1 partition derived from the original wishart matrix; well the having a Wishart distribution, Wishart k n sigma 1 1; similarly, one also will be having A 2 2; the partitioning the second block partitioning, that would be having a Wishart distribution on m minus k dimensions with m as the degrees of freedom, and sigma 2 2 as the associated variance, covariance matrix; this can also be proved.

Let me, write the third result: which is for the wishart partitioning; we also have the third result: as if we look at this A 1 2 given A 2 2. This will be having a matrix normal distribution. This will follow a matrix normal distribution n with the following parameters: sigma 1 2, sigma 2 2, inverse A 2 2, as the main matrix, and the associated variance, covariance matrix would be given by sigma 1 1 dot 2 to make a product A 2 2; where this sigma 1 1 dot 2 is the usual sigma 1 1 dot 2 matrix. That is, it is sigma 1 1, sigma 1 2, sigma 2 2, inverse multiplied by sigma 2 1.

So, the conditional distribution of A 1 2, that partitioning of the Wishart given A 2 2; this now follows, a matrix normal distribution; note that this is a rectangular matrix. So, no question of having this conditional distribution to be having Wishart distribution, because Wishart distribution is associated with the symmetric matrix, and the third or rather the forth, and the last result: concerning the partitioning of the Wishart distribution is the following: if we define A 1 1 dot 2 as in the similar way as sigma 1 1 dot 2, which is A 1 1 minus A 1 2, A 2 2 inverse, A 2 1. This now, will follow a Wishart distribution the order would be same as A 1 1 or the order of this. So, that is k; this would follow Wishart distribution k on degrees of freedom as n minus m plus k, and the associated variance, covariance matrix as the sigma 1 1 dot 2, and is independently distributed of A 1 1, and A 2 2. So, these are some fundamental results concerning partitioning of the Wishart distribution.

(Refer Slide Time: 34:35)

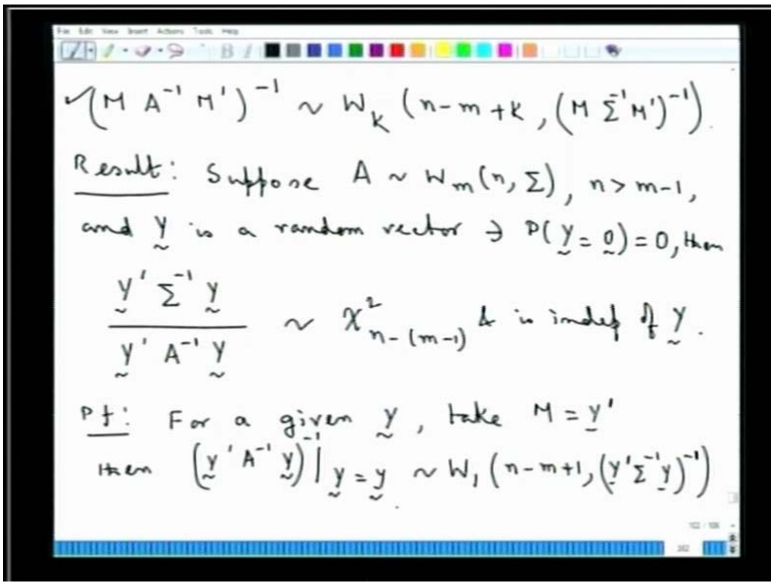


Now, the next important concept: that we are going to introduce since an inverted wishart; what we call as by an inverted Wishart the definition of inverted Wishart is same to (()), and would eventually lead us to hotelling T squared distribution. So, if we have A to follow, a Wishart distribution, Wishart m n sigma. Then, the inverts of this matrix inverse of this random matrix, that is A inverse is said to follow an inverted Wishart distribution.

This inverted Wishart actually, would also lead us to an unbiased estimator of the sigma inverse; which we are going to see shortly. Now, an important property an important property of inverted Wishart is the following: inverted wishart is the following that let me, write it completely suppose A follows a wishart m n sigma; sigma is positive definite m is say k by m matrix of constants, matrix of constants and rank of m is full.

So, suppose we have this particular setup; then for of course, n greater than m minus k. We will have the following distribution; that M A inverse M transpose whole inverse, this would follow a Wishart distribution k on n minus m plus k, and the associated variance, covariance matrix as M sigma inverse M transpose whole inverse.

(Refer Slide Time: 37:41)



So, what is this result basically telling us that, if we have A to have a Wishart distribution, and if we have m a matrix of constants k by m order of rank k. That is, it is a full row rank, and for n greater than n m minus k; that is what is require in order to ensure that degrees of freedom of this Wishart distribution. This is basically, n minus m minus k. So, we ensure that this is greater than zero. So, we have the degrees of freedom strictly, greater than 0. Then, we have this M A transpose M inverse whole inverse will be having a Wishart distribution n minus m plus k, and the associated variance, covariance matrix as M sigma inverse M transpose whole inverse.

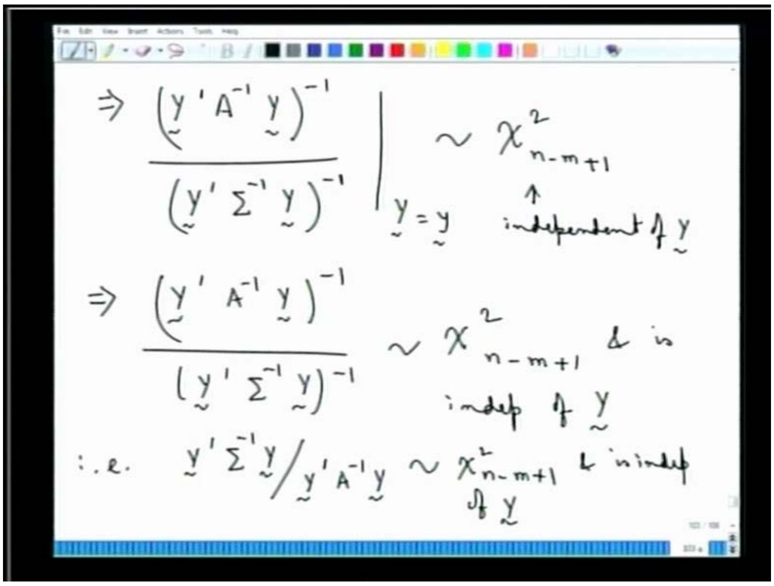
Now, using this result of this Wishart distribution, we have the following important result: suppose, we have got A to be a Wishart m, n sigma, n is greater than zero, positive integer which we take n to be greater than m minus one, and sigma of course, is positive definite, and Y is a random vector y is a random vector such that probability Y equal to y equal to null vector is 0. Then we will have this Y transpose sigma inverse Y; this divided by Y transpose A inverse Y; this will now, be having a chi square distribution central on n minus m minus 1 degrees of freedom, and is independent of this Y vector. right

So, this reminds us of a similar result: that we had proved today; which actually was without the inverse, it was Y transpose A Y by Y transpose sigma Y. It was shown to have a chi square central distribution on n degrees of freedom. Now, instead of working with the Wishart distribution; if we are now working with an inverted wishart distribution; then this is basically, the result which tally's with the result that we are previous previously obtained. So, this result is simple; actually, it basically uses this fundamental result concerning an inverted Wishart distribution; which we said to have an important property of an inverted Wishart distribution. So, the proof of this goes along the same line as proof of the result for the Wishart distribution.

So, for A given Y, if we take this M, and this result as Y transpose then, what we will be having that this Y transpose sigma Y transpose A inverse Y. It is basically, we are using this particular result which with M equal to Y transpose. So, this for a given Y equal to say small y the distribution of the inverse of this quantity will follow a Wishart distribution with what dimensionality.

Now, k is the dimensionality associated with the M matrix; now Y transpose is taking its place, and hence we will have this the dimensionality of the wishart distribution to be equal to one, and what is going to be the degrees of freedom; it is going to be n minus m minus k. So, that is n minus m plus k. So, that this is n minus m plus one. So, that is the degrees of freedom of the Wishart distribution, and what is associated made in (( )) that is going to be equal to Y transpose sigma inverse Y, then this is whole inverse.

(Refer Slide Time: 43:01)



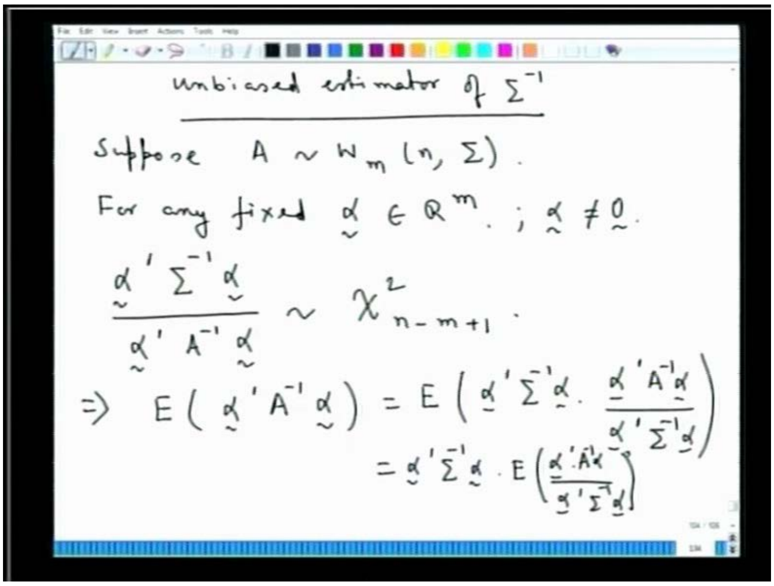
So, this result is used in order to get this particular form here. So, that what we now have now we have already, seen that what happens to a Wishart distribution on one degrees of freedom. So, this would imply that this Y transpose A inverse Y inverse of this; that that was having the Wishart distribution of one degrees of freedom. This divided by the associated variance, which is Y transpose, sigma inverse, Y whole inverse, this for a given Y equal to small y. This would follow, what a chi square distribution on n minus m plus 1 degrees of freedom; why because of the simple fact that we have got this the distribution of this given Y equal to y has got a Wishart distribution n minus n plus one, and this as associated variance term, and hence if we divide this term here by the corresponding variance. We are going to have a chi square random variate on the degrees of freedom associated with the degrees of freedom of the Wishart distribution.

Now, this thus is the conditional; this chi square n minus m plus 1 is thus the conditional distribution of this random variable. Here, Y transpose A inverse Y inverse divided by Y transpose sigma inverse Y. This is basically having a conditional distribution condition by Y equal to small y is having this. Now, this distribution is what we observe is independent of this conditioning variable y, because whatever be the capital y being fixed at small y that this distribution is going to be a chi square central on n minus m plus 1 degrees of freedom; thus the conditional distribution of this expression given Y equal to y would be same as that of the unconditional distribution.

So, this would imply that, this Y transpose A inverse Y; this inverse that divided by our Y transpose sigma inverse Y. This is the unconditional distribution; this is going to also have a chi square on n minus m plus 1 degrees of freedom, and this is going to be the distribution of this random variable; since, it is having the unconditional distribution same as that of the conditional distribution, and this thus would be independent of this conditioning variable conditioning random vector which is Y, and is independent of this conditioning random vector; which is Y that is, this is just the inverse of that. So, that we have this Y transpose sigma inverse Y this divided by this Y transpose A inverse Y this to have a chi square n minus m plus one, and is independent of this Y vector, and that was actually the result; which we try to prove that sigma Y transpose sigma inverse Y divided by Y transpose A inverse Y, this has got this desired distribution.

Now, we look at an unbiased estimator of sigma square; that is associated with the inverted Wishart distribution. Let me, look at an unbiased estimator, this is all associated with an inverted Wishart distribution, and what we will be saying is that inverted Wishart also is going to be used in order to derive to the hotelling T squared distribution unbiased estimator of this sigma inverse.

(Refer Slide Time: 46:43)

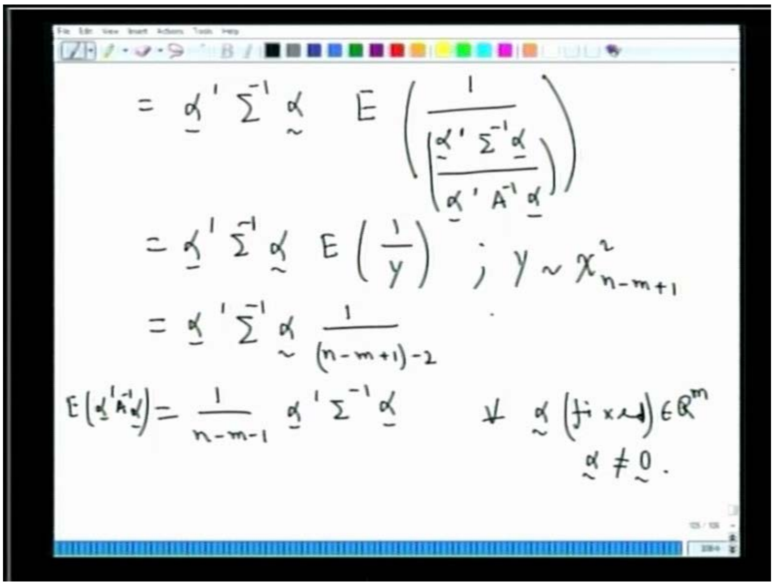


Now, in order to derive this unbiased estimator of sigma inverse; suppose, we have got a Wishart distribution suppose A has got a Wishart distribution, Wishart m n sigma; then for any fixed vector alpha belonging to R to the power m of course, we take this alpha vector to be not equal to a null vector for this fixed alpha vector; what we can write is the following: that alpha prime sigma inverse alpha. This divided by alpha prime A inverse alpha; what is going to be the distribution of this by the previous result; that that is what we had proved, if we take alpha if we take alpha to be equal to Y in the previous result with alpha degenerate at a particular point with alpha not equal to zero; ensuring that this is not equal to zero. So, that this will have a central chi square on n minus m plus 1 degrees of freedom; that is why the previous result with Y degenerate at this alpha point with alpha not equal to A null vector satisfying the conditions of the previous result.

Now, this would imply that, if we now look at expectation of alpha prime A inverse alpha; now this, what we are doing for a general Wishart distribution; we will use that in order to get this unbiased estimator of sigma inverse expectation of alpha prime A inverse alpha. Let me, write that in the following way, that this is alpha prime sigma inverse alpha that multiplied by alpha prime A inverse alpha that divided by alpha prime sigma inverse alpha.

Note that, this is this part the first part is constant. So, that we will have this written as alpha prime sigma inverse alpha that multiplied by expectation of this alpha prime A inverse alpha that divided by alpha prime sigma inverse alpha. right

(Refer Slide Time: 49:39)

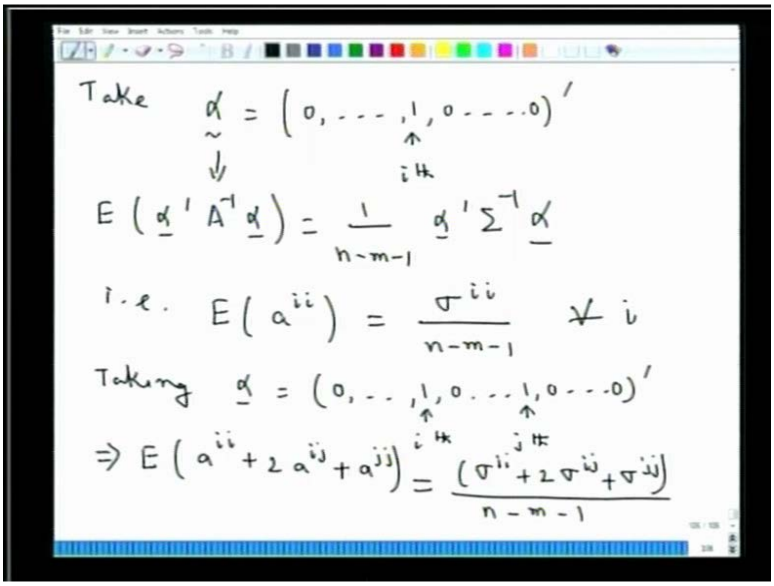


Now, what is this quantity equal to (( )). So, that we can just find that, this expression is equal to alpha prime sigma inverse alpha, and then expectation of 1 upon alpha prime sigma inverse alpha; that divided by alpha prime A inverse alpha, why have written in this particular form, because we know that this has got a chi square distribution. So, let me just write this simply as alpha prime sigma inverse alpha into expectation of 1 upon Y say, where this Y follows a chi square distribution on n minus m plus 1 degrees of freedom.

Now, if Y has got a chi square on n minus m plus 1 degrees of freedom; we know that expectation of 1 upon Y would be 1 upon degrees of freedom minus 2. So, that this expression straight away rewrite it using univariate distribution theory; that this is going to be equal to 1 upon the degrees of freedom of this that is n minus m plus one, this minus 2; Y is this true, if Y follows a chi square on n degrees of freedom. Then, expectation of 1 upon Y is 1 upon n minus 2. So, that we have this as 1 upon n minus m minus 1 into alpha prime sigma inverse of alpha. So, that our expectation of alpha prime A inverse alpha is equal to this term.

Now, this is true for every alpha every fixed alpha for every fixed alpha belonging to R to the power m alpha of course, not equal to this null vector; otherwise there will be problem in defining this particular random variable. Here, now let me use this particular term; that is what we have now? If we take some special choices of this alpha vector.

(Refer Slide Time: 51:56)



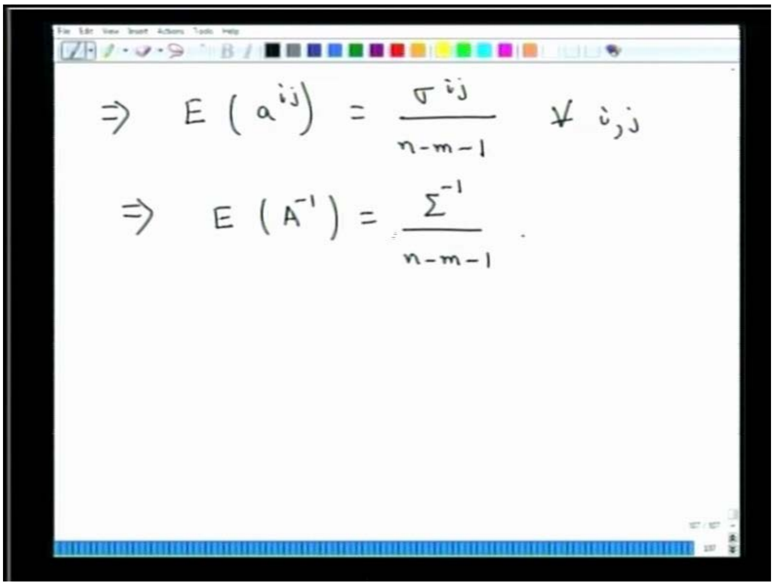
They will have the desired result take this alpha vector to be 0 at all the positions, except the i th position; say, if we take this as one, and at all other locations. If we have zero, then what is this expectation of alpha prime A inverse alpha; this we know is equal to n minus m minus 1 into expectation of, I am sorry this is expectation we have already kept; it this is alpha prime sigma inverse alpha.

So, with this alpha in this result, we will have here expectation of a upper i i; say a upper i i is the i i th element of this A inverse matrix; that would be given by this alpha prime sigma inverse alpha would be sigma i i, sigma upper i i is the i i th element of this sigma inverse matrix. That divided by n minus m minus one; this is going to be true for every i. So, we have obtained that expectation of a i i; a upper i i is equal to sigma i i by this now further take alpha equal to another choice taking alpha to be a vector; which is having zero at all the positions except the i th, and the j th position. So, this is the i th position, and this say is the j th position, and zero at all other positions.

One again using this positions result what the alpha prime A inverse alpha would lead us to this would imply that, expectation of a upper i i plus, 2 times a upper i j, this plus a upper j j. So, these are all the elements corresponding to A inverse matrix; this would be equal to from the right hand side, alpha prime sigma inverse alpha is going to be sigma upper i i, this plus 2 times sigma upper i j, this plus sigma upper j j; that is coming from alpha prime sigma inverse alpha, and this is n minus m minus one.

Now, we already have a expectation of each of these a i i to be equal to sigma i i by n minus m minus one, and hence this can be replaced by expectation of a i i by n minus a into m minus one.

(Refer Slide Time: 54:50)



So, this would further imply this would imply that expectation of a upper i j, that would be given by sigma upper i j that divided by n minus m minus 1. So, what have we proved now this is going to true for every i, and j. So, that this is the off diagonal entries, these are the diagonal entries this is true for every i equal to 1 to up to m, and we have this for every i j equal to 1 to up to m. So, this would imply that expectation of A inverse. So, every element has got it is entries there, so expectation of A would be given by sigma inverse by n minus m minus 1.

So, we will in the next lecture, we will use this particular result in order to get to an unbiased estimator of sigma inverse, when we have a random sampling from a multivariate normal distribution from this particular result. Thank you.