Applied Multivariate Analysis

Prof. Amit Mitra

Prof. Sharmishtha Mitra

Department of Mathematics and Statistics

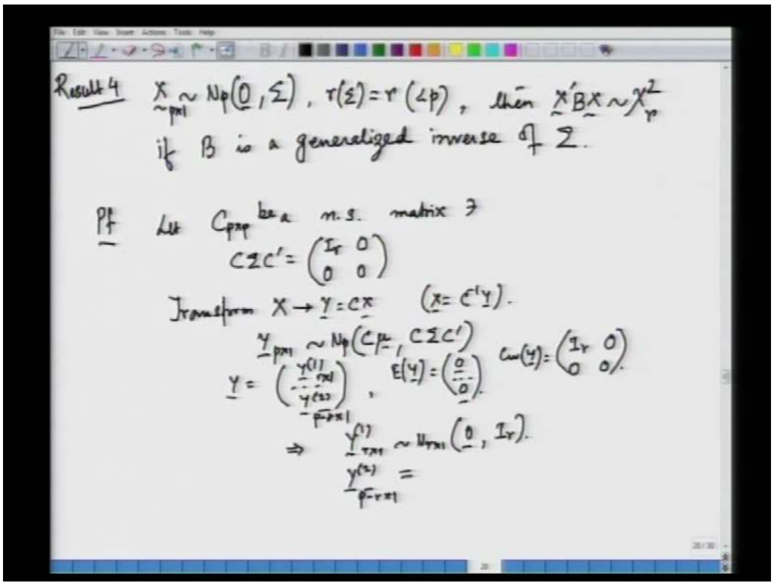
Indian Institute of Technology, Kanpur

Lecture No. # 05

Multivariate Normal Distribution - III

We are beginning this session with a result, where we are going to give up the restriction of positive definiteness of the covariance matrix sigma.

(Refer Slide Time: 00:29)



So, our first result in the series is there be a continuing, so this is a result four, and we take that X is following a p variate, multivariate normal distribution with the mean vector mu and covariance matrix sigma. As mentioned, we have rank of sigma now this is p-dimensional, so this is p-dimensional square matrix, but we take the rank of sigma as r, which is strictly less than p. So, this is the important relaxation that we are doing; this is a situation, then the quadratic form X transpose B X say this is going to follow. Let us take a simplification here, and then can this can easily extend to the non-null vector also, but let us state it in this form.

So, that we consider simply X transpose instead of taking X minus mu transpose, and this as result follows the standard chi square distribution with r degrees of freedom, which is a rank of the sigma matrix, and the condition on the associated matrix B is that if B is a generalized inverse of sigma. So, you can see we have situation of single multivariate normal distribution here, sigma is positive semi definite, as we had said in some of in one of earlier session that is the multivariate normal distribution is singular multivariate normal distribution, where the pdf does not exist, because sigma inverse does not exist. And so obviously sigma inverse does not exist. So, what we are working with is a generalized inverse of sigma.

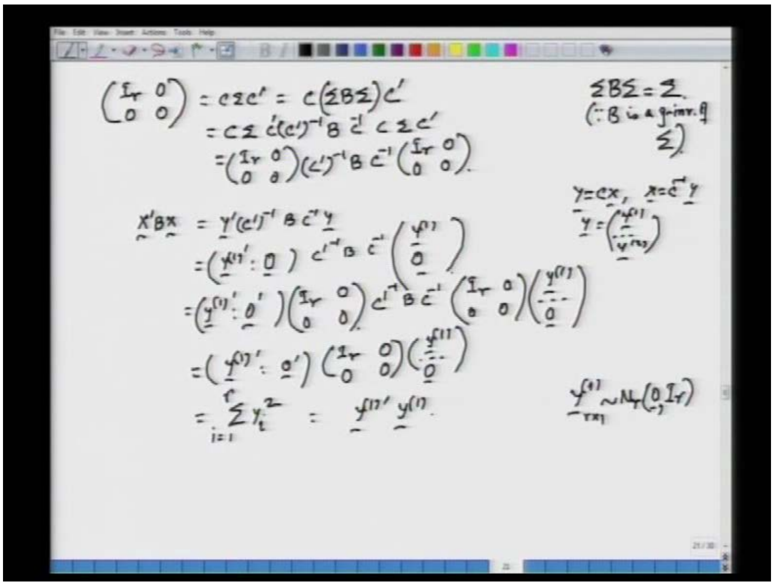
So, let us go to prove of this. So, we start with the proof of the result and in the very beginning we consider some matrix p, let C a square matrix of same dimensions as sigma a nonsingular matrix such that, we have C sigma C transpose is the identity matrix of dimension are rest of the blocks are null matrixes basically what we are doing is we are considering a rank factorization and we are writing in this form. So, we have say sigma C transpose I r rest of them as null blocks and then we consider the transformation. So, now we transform from X we are going to Y which is C X, so that we have X is nothing.

But C inverse Y no problem in it, because C is nonsingular and then what we C is Y are earlier result we have Y which is also a p dimensional random vector this follows a multivariate normal distribution p dimensional with mean vector C mu and covariance matrix is say sigma C transpose it is, just the matrix that we have considered here we have partition of the covariance matrix as well in the form of I r null null null. So, basically now if I partition the Y matrix accordingly. So, that the partitioned vectors are now giving me this covariance matrix. So, that I have Y into Y 1 and Y 2 what I do is I can clearly see that the covariance of the first part first component is in here in the first block of this C sigma C transpose.

So, basically this has been partition into r and p minus r components. So, this is the first one is r dimensional vector the next one is p minus r the residual is coming here and so that I have well expectation of Y is nothing but I consider this as the life started with Xfollowing normal p with mean null vectors. So obviously, this is also both of them a null mu 1 and mu 2 what is happening is covariance of Y and this is nothing but I r.

So, this implies that the first component the first sub vector of dimension r this is a multivariate normal distribution with mean vector is null and the covariance matrix is the identity matrix of dimension r and what is happening to the second component well this component a p minus r dimensional vector it has mean the null vector and the covariance matrix is the null matrix. So, what is happening is in essential I have this variable this Y 2 random vector is actually a degenerate random vector and obviously degenerate act the origin at the point zero.

(Refer Slide Time: 06:22)



So, this is equal to the null vector with probability one p continue with this so let us have a look at the C sigma that is the covariance matrix of the transformed variable Y C sigma C prime which we are writing in this form now. And this is as I said this is nothing but the covariance matrix of the new variable Y now here we are using the most common and most popular definition of sigma of generalized inverse of a matrix. So, what we had started with we had set that we is a generalized inverse of sigma.

So obviously, what we have is sigma B sigma is nothing but sigma since B is a g inverse of sigma. So, I can replace this sigma by the sigma B sigma and I have a C transpose at the end. Now this is a little bit of manipulation or little bit of trick here, I just import C transpose C transpose inverse here, after sigma then I write B and then again I what I bring in forcibly is C C inverse and then once again I write sigma C transpose now there is no problem and doing so because I have C as nonsingular.

So, what I have done is a basically imported this C transpose C transpose inverse which is an identity matrix. So no problem I can insert it anywhere, I like similarly I have in sorted this C C transpose also which is again and identity matrix of dimension p. So, no problem now with this first I am combining the first part the first matrices giving me I r and then the rest of it what remains here is C transpose inverse B. Again I have what I will do here is instead of writing. So, that I can very conveniently again have a C sigma C transpose I will I write this is in this form. So, this is C transpose C inverse C right no problem in it whether it C C inverse C inverse C.

So, I have a C inverse here and then I have C sigma C transposes giving me again this matrix I will have this stage and then I concentrate on the quadratic form that we had take a now the Xtranspose B Xand let us see what it is. Well, what was my transformation the transformation was Y is C X basically X is nothing but C inverse Y and I use it here to get Y transpose C transpose inverse B C inverse y. And now I use the partition form of the random vector Y and write this as recall that Y was partition into Y 1 and Y 2 and we have transpose here. So, that this is Y 1 transpose of maintained column wise this is Y 2 transpose and then I have C transpose inverse B C inverse again I am writing Y 1 and Y 2.

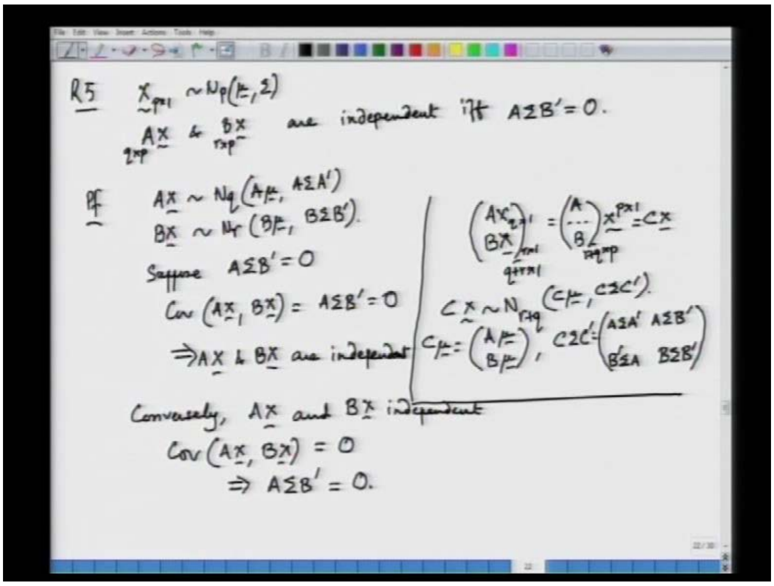
Now, what I do is now after this I insert the matrix I r null null null. Now this Y 2 we have already seen that this is nothing but this is random vector is degenerated zero with probability one. So, I can straight away right this is null vector. Similarly, here also let me replace Y 2 by the null vector, here also I am going to do the same thing this system null vector what I am doing is a I am bringing in this I r null null and null matrix here it is making no difference. Because this is practical acting as the identity matrix as per as this vector is consider and then I write this the rest of it sigma transpose inverse B C inverse and then again bring in this same matrix and writing the vector at the end.

Put a transpose here the make of difference. So, this is Y 1 transpose and we have this and this has been shown to be equal to this matrix itself. So, simply I have this is getting replaced by the single matrix which has been proved in the earlier step and then I have is coming up which shows that this quadratic form is nothing but the some of Y I square I from 1 to r only.

So, this is basically Y 1 transpose Y 1, now with Y 1 random vector are dimensional following multivariate normal with mean as the null vector and the covariance as I r these are independent and I can have the sum of square following a central chi square distribution with are degrees of freedom. So, I have X transpose B X sigma the covariance matrix of X was not positive definite, but its just positive semi definite with rank equal to r which was strictly less than the dimension p and what was the only thing that we assumed here is something about this associated matrix B we have take in this B as one generalized inverse of sigma.

So, this proof is a completed this step. Now we move to our next part of the results, the next group of results now till now we have considering either a linear form or a quadratic form in isolation and trying to show what the what distribution is now what we are going to do is we are going to consider two of them. There is two linear forms a linear and a quadratic form or two quadratic forms and try to check there whether they are independently distributed or not are basically talk about the condition then they will be independently distributed or not. This sort of independence the establishing independence will be useful when we later go to the inference part and there we will see that to talk about to establish the test statistics the distribution of the test statistics independence of this quadratic forms are really important.

(Refer Slide Time: 14:10)



So, next result is result five we start with the simplest one. We say that X is multivariate normal mu and covariance matrix sigma we have two linear form say A X and B x. So, A and B did not be square matrixes, this kind A can be q by p rectangular matrix where is B can be and r by p rectangular matrix and I will say that the two linear forms this is not y, but X the same variable X are independent. If and only if A sigma B transpose is a null matrix group is really simple. So, we first have look we have already establish this preliminary things. So, which is that A X is itself multivariate normal q dimensional with A mu A sigma A transpose as the covariance matrix B X the other linear form of X is another multivariate normal with mean B mu and covariance B sigma B transpose and I am assuming suppose A sigma B transpose is in fact a null matrix.

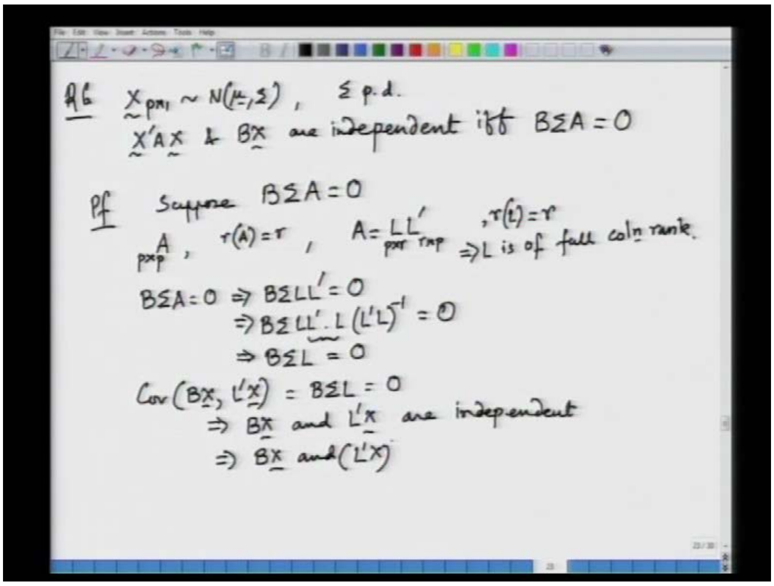
Now, what is covariance between A X and B x? Well, this is certainly A sigma B transpose which has been assume to be equal to one null matrix. Now, what is happening here let us see if I augmented A X and B X what I am getting well I am getting a dimensional vector here and this is sorry this is q we have take in a as A X is q dimensional and B X is actually r dimensional. So, the augmented one is q plus r dimensional and this can be actually written the form as A augmented B x.

So, this is I have a r plus q dimension p and this is p dimensional. So, let us call this as some C x. So, this is the C X is now again going to follow a multivariate normal distribution with r plus q and this mean is say C mu and the variance the covariance matrix C sigma C transpose, but then what is C mu what is the form of C mu well it is nothing but A mu augmented B mu. And what is the form of C sigma C transpose it is A sigma and A transpose in the first block then it is A sigma B transpose of this here.

So, it is B sigma B transpose a sigma and then we have B sigma B transpose now what we have here by here assumption is this half diagonal blocks are null matrixes. So, that we have by our very special feature of multivariate normal we say that this A X and B X is independent. So, this is something which we have used for the proof and then we have the converse part that is really simple. We have A X and B X are independent whether there multivariate normal I do not care about distribution I will always have covariance between A X and B X in that case as this null matrix and this implies because this is covariance nothing but a sigma B transpose this null.

So, this is simple proof of the independence between two linear forms of multivariate normal random vector A X and B X, now what we have going to consider is one linear form and one quadratic form and see what is the condition for the independence between those two.

(Refer Slide Time: 19:22)



So, our next result is I take in very general set up that this is following normal mu sigma with we are again that two our usual important assumption, that sigma is positive definite and what I have is one quadratic form X transpose A X and one linear form which is B X simply they are independent. If and only if B sigma A B is the associated matrix of the linear form and A is the associated matrix of the quadratic form this is null. So, for the first part of the proof that is saying let us start with the sufficiency part and we assume suppose B sigma A is in fact null. So, what we do is we are considering a rank factorization of A.

Now, A obviously is p by p because extra transpose A X is defined now there is no need to consider of full rank. So, I can have rank of A as r. So, I can consider a rank factorization of a saying to L and L transpose where L is half dimension p by r with rank of L equal to rank of A which is r. So, basically it tells me that L is half full column rank we are going to use this later. So, this is nothing but rank factorization of this matrix A which is appearing in the quadratic form of X.

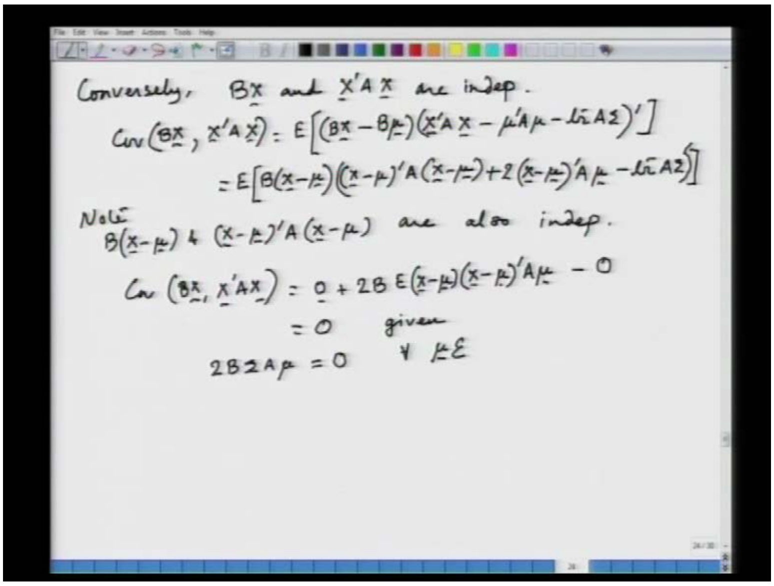
So, now we have B sigma A given that this is null this only implies that a now we are going to use that A is nothing but L L transpose. So, this is now so what I do now is I bring in I already have L L transpose X try to L L transpose L inverse this side still remains a null matrix no problem in doing this what I have brought in is L transpose L inverse which exist, because as we have already said that L is a full column rank not only this matrix L transpose L is r dimensional with its rank of L transpose L is nothing but equal to rank of L which is equal to r.

So, L transpose L is of full rank and its inverses so I bring it in here the other side is still null. So, X tries this part and let us sees why we do this how it is going to help us it is almost immediate, because I see that I can get read of these four matrices. So, that this is basically two matrices here combining L transpose L and L transpose L inverse and I can write that B sigma L is the null matrix. What is the advantage if I write B sigma L is null I already had B sigma A is null. Now I am establishing B sigma L is null what I can do is I can use my earlier result and saying that case that covariance between two linear forms B X and L transpose X which is nothing but B sigma L and which we have just now shown to be equal to a null matrix this is equal to zero and by are earlier result I immediately have B X and L transpose X are independent.

If B X and L transpose X are independent I can might as well say that B X and L transpose X suppose this is some Y if I have X and Y are independent I can variable say that X and Y transpose Y are also independent. So, that is exactly what I am doing here I say that B X and L transpose X transpose L transpose X they are also independent, but what is this after all this is nothing but our quadratic form X transpose A X they are independent. What are the distribution where like and see that this B X also has a multivariate normal distribution what about X transpose A X it has a non-central chi square distribution that does not matter.

We can start with the general mu if it is null vector, then we will have a central chi square independence between central chi square and multivariate normal here it is incidentally on non-center chi square. So, this is one part of the proof where we have assumed to this B sigma A is null matrix.

(Refer Slide Time: 24:55)



Let us go to the other part, so conversely we are assuming B X and X transpose A X linear and the quadratic forms they are independent. So, I need not care about the distribution of these I simply consider the covariance between this one is random vector, this one is a random variable, and I consider covariance between B X and X transpose A X, I use the very definition of covariance between two random variables, and this is nothing but expectation of first one B X minus its expectation what it is B mu. And then it is the second one minus its expectation where we have already seen what is the expectation of this quadratic form this is nothing but mu transpose A mu minus trace of a sigma with the transpose.

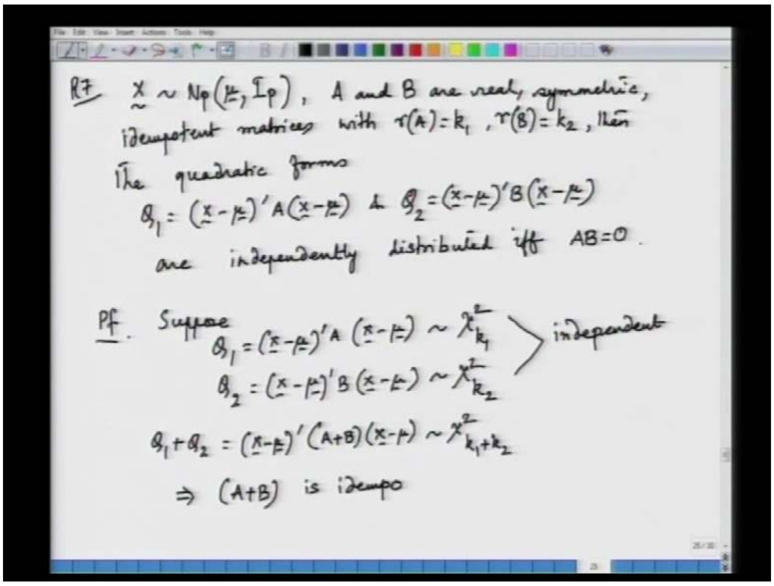
So, this is it. So, now, what we are going to do is we considered is this is we coming out and we have B with X minus mu. So, just like we have this location shift here it is now X minus mu with try to do this same thing in this part of also in the other part. So, that now I write it forcibly as X minus mu transpose A X minus mu when I do so well I have already included the X transpose A X and mu transpose A mu, but they are some extra cross product terms that I have introduce and I must make adjustment for this and hence that adjustment is giving me nothing but two X minus mu transpose with A mu and I have the other term that is constant term remaining which is trace of X sigma this with the transpose.

So, now what we are going to uses we have been given B X and X transpose A X are independent. Now which also implies that we have since B X and X transpose A X are independent it is same as saying that B X minus mu and X minus mu transpose A X minus mu note are also independent and then what do I get from that covariance term I have I am considering expectation of this B X minus mu with this product this is the first thing joint expectation that I am considering.

Now, since this are independent well using the result that expectation of X Y is expectation X times expectation Y I am doing just that and note that as soon as I take expectation of the first part I end up with a null vector. So, I have this as null here then what is happening in this part. In the second part, well I have twice plus sign then the matrix of constant coming out I have expectation X minus mu with X minus mu transpose and I have another A mu and what is happening with this part will this is constant part totally and as soon as I take expectation this is the only non-stochastic this is the only stochastic part and as I take expectation on this is going to be null.

So, this part is null and I have I had started with that these two are independent this is equal to zero this is given. So, what I have now is basically two B this is what this nothing but the covariance of x. So, this is basically it is p sigma A mu and this is equal to the null vector for all mu because I given condition that B X and X transpose A X are in independent this is null for all mu in R p which will imply that associated matrix has to be a null matrix. So, this is what we wanted to prove that if the linear form and the quadratic form they are independent what I have is the associated matrices relationship involving then in this way. So, p is the associated matrix of linear form, sigma is a covariance matrix of the random vector, X and A is the associated matrix of the quadratic form and they together is null matrix.

(Refer Slide Time: 30:57)



So, this complete step proof a next one is about the independence of two quadratic forms. So, we had started with the independence of two linear forms, next one was one linear and one quadratic and now we come to the independence between two quadratic forms. So, this is result seven we state this x following normal p start with the setup of mu I p and very conveniently use this setup in the general setup where we have x following a multivariate normal distribution with mean mu and the general covariance matrix sigma.

So, first establish it first sigma equal to the identity matrix the uncorrelated case. So, now then we have A and B are real symmetric idempotent matrices with some rank of A is k 1 and rank of B is k 2. Then the quadratic forms Q 1 say the first one is X minus mu a deliberately taking the siltation shift to end up with central chi square distribution is the first one is X minus mu transpose A the first matrix coming into the picture X transpose X minus mu transpose A X minus mu and Q 2 is the second one X minus mu transpose B X minus mu are independently distributed.

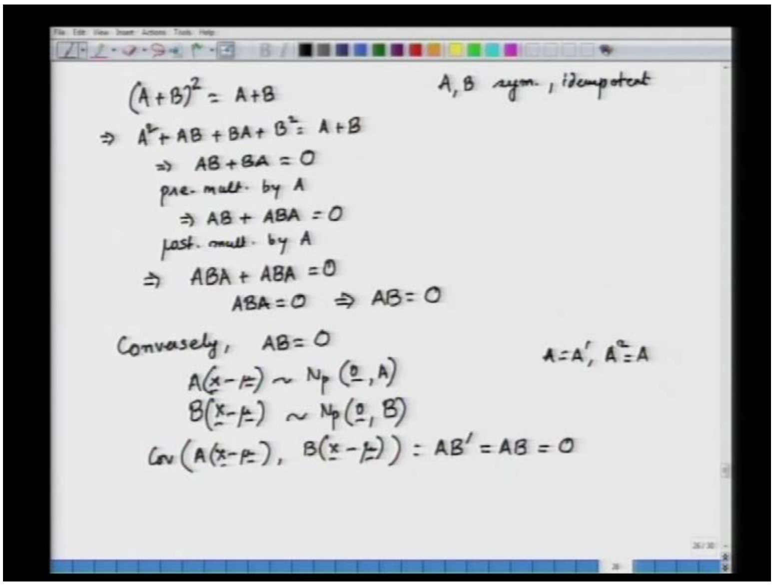
If and only if the product of the two matrices A B is a null matrix so, for the proof of this result we start with let us start with the necessity part that is the only if part and we suppose that the quadratic form incidentally with the given condition what are the distribution of this quadratic forms Q 1 which is X minus mu transpose A X minus mu.

This will follow a central chi square with rank of A as its degrees of freedom and Q 2 this is X minus mu transpose B X minus mu this is also following a central chi square with rank of B as its degrees A freedom Q 1 and Q 2 independent. So, we start with this assumption and then complete the proof.

So, we have to independent slip distributed chi square variables. So, their some by the additive property of chi square distribution of independent chi square distribution this is also going to be a chi square variable. So, that we have Q 1 plus Q 2 which is nothing but X minus mu transpose and we have the matrix matrices a adding A plus B X minus mu this also follows chi square with k 1 plus k 2 as the degrees of freedom now from are earlier result we know that, if this is a chi square then the because that result was an if an only if result.

So, this is going to imply we have from here that A plus B is an idempotent matrix. Not only that it is we also had seen their that rank of A plus B is actually equal to the degrees of freedom that is rank of A plus rank of B in the special situation. So, A plus B is idempotent now we use this idempotentancY of a plus b to establish what we have to show is that A B is a null matrix.

(Refer Slide Time: 35:53)



So, if A plus B is idempotent from the definition of idempotentancy, I have A plus B square is A plus B which means that A square plus A B plus B A plus B square this is A plus B. Now, note that there are many properties of A and B they are symmetric and they are idempotent. So, I have A square is A and B B square is B. So, this is cancelling out with these matrices and I am getting A B plus B A is a null matrix the addition is a null matrix. So, what it does is this is just one way of proving it. So, what I do is a pre multiply by A and this gives me A square B which is again A B because A is idempotent plus A B A this is null obviously and the next step what I do is I post multiply by A again.

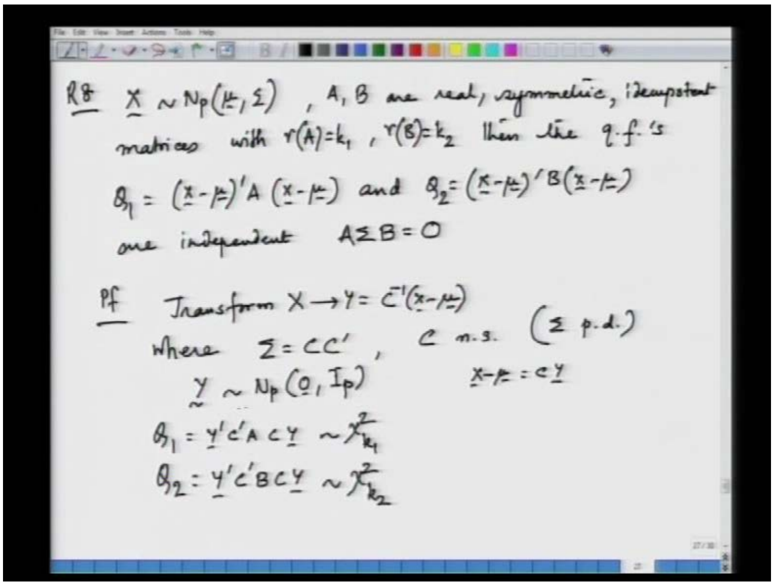
So, what I get is this implies what I have is A B A plus A B A square which is A B A and basically what I end up with is A B A is a null matrix. And using this step over here, this is going to imply that A B is also a null matrix because there some is a null matrix. So, I prove the necessary part of it and then the sufficiency part conversely here I am going to assume that A B is an in fact a null matrix what I consider at this stage is A times X minus mu this is now we null this things this is p variate multivariate normal with mean vector null because I already have made this shift here and what is the covariance matrix is the covariance matrixes nothing but A A transpose because the covariance matrix in this set up is actually the identity matrix.

Now, A A transpose is actually a A times A, because A is A transpose and A square is nothing but A so I end up with a p variate normal zero A. Similarly B X minus mu also follows same dimension null vector and B as the covariance matrix now what I have is if I consider covariance between these two that is a X minus mu and B X minus mu what I get is, because that covariance matrix is nothing but an identity matrix.

I have this as A B transpose, but B Again being symmetric this is A B. And now this has been given to be equal to zero again by our earlier argument or earlier result also we are proved we have these A X minus mu and B X minus mu these are independent, which immediately gives us what I do is I again take the transpose of this and multiply it with itself that is if I X and Y independent I am considering that X transpose X and Y transpose Y are also independent. So, by that remain is I have A X minus mu this transpose A of X minus mu and B X minus mu transpose B times X minus mu these are independent and this is giving me X minus mu transpose again get A transpose A which is A square which is equal to A.

So, I have a X minus mu and X minus mu transpose B X minus mu these are in fact the two quadratic forms Q 1 and Q 2 these independent if A B is a null matrix. So, this completes the other part of the proof also as I had say we had started with the simpler situation where we are a assume the covariance matrix to be identity matrix now we are going to go to the more general set up where we have the usual positive definite sigma matrix, but we are going to directly use of proof of this result to establish the result in that situation.

(Refer Slide Time: 41:50)



We have just a small modification which is very obvious. So, this is now result eight this says now I have X of a p variate multivariate normal with mu sigma and A B as before are real symmetric idempotent matrices with rank of A as k 1, rank of B as k 2, then the same quadratic forms then the quadratic forms Q 1 first one was with the A x minus mu transpose A X minus mu and the next one these are independent.

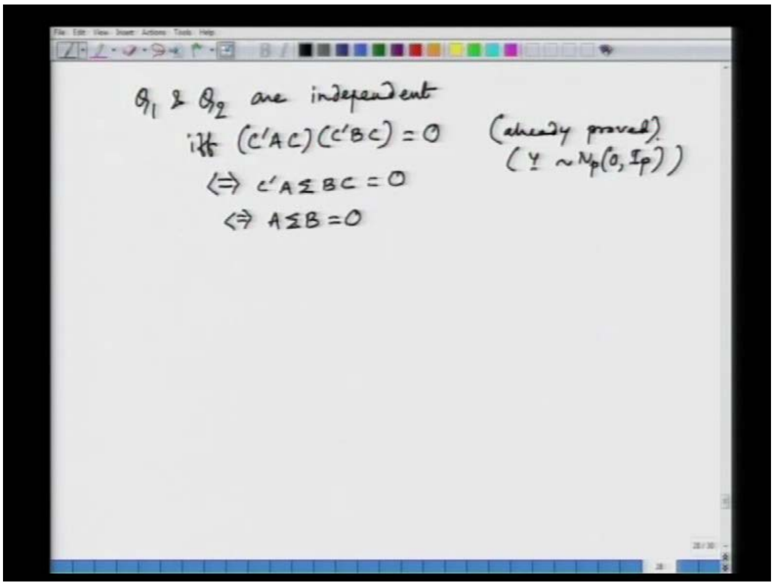
You can guess that earlier were we had just the product of A B null matrix here we have a sigma coming into the picture. So, that is sigma gets inside and we have A sigma B is a null matrix. So, we are going to prove this one with the help of the case sigma equal to I that we have to prove just now for that what we need to do is, just us simple transformation and I take transformed I take some Y says. So, I am transforming from X to Y and I am saying that this is some C inverse X minus mu what sort of C is going to help me as where sigma this has been decomposed into C C transpose and C is nonsingular will C is nonsingular means well I have obviously, take in sigma as positive definite and not positive semi definite.

So, if this is then what is the distribution of Y will then Y in that case follows a p variate normal with mean vector null and identity in matrix of order p as the covariance matrix I have I also have X minus mu this is basically C y. So, that the first quadratic form Q 1 in terms of Y and C is nothing but well I am using that X minus mu is C y. So, what I have is Y transpose C transpose then A and for X minus mu C y.

So, this is nothing but Q 1 and this is under the given set up where A is idempotent with rank equal to k 1 this is nothing but the central chi square with k 1 degrees a freedom which incidentally is also rank of this matrix C transpose A c, because C being a nonsingular matrix does not alter rank of A when A is getting pre and post multiplied by C transpose and c.

So, this is chi square central chi square with k 1 degrees of freedom what about Q 2 this is Y transpose C transpose B C Y and this follows gain this, because this is Q 2 under the given set up where b is idempotent with rank equal to k 2. So, this has two follow a central chi square with k 2 degrees of freedom and I have in the situation where I have this sigma as identity matrix. So, now which I have exactly the same situation now because more over mu is also simplified to a null vector and the covariance matrix is in fact, and identity matrix the case I have show proved earlier.

(Refer Slide Time: 46:27)



So, by our earlier result Q 1 and Q 2 are independent. If and only if product of the associated matrices that is a null matrix. So, the product of the associated matrices the first one is C transpose A C and the next one is C transpose B C this is null. Now this result incidentally already proved, because here the random vector Y has a set up with the identity matrix as the covariance matrix and this has been proved for this case.

So, this is so what remains to prove is that this equal to null is equivalent to what we want to show to be null that is A sigma B is null. So, this is a null matrix means that I have C transpose A C C transposes nothing but sigma, then I have B and C, this implies that, I can since C is nonsingular I can pre multiply by C transpose inverse, post multiply by C inverse to get A sigma B is… And each of these steps can be trace back, so this is a null matrix is in fact equivalent to saying that A sigma B is a null matrix, because C inverse exists, and this is what we wanted to show.

So, we are going to stop with this result in this type in the series of this type of results, the next topic that we are going to take up is sampling distribution from multivariate normal distribution. So, that I can get I can get sufficient statistics for the parameter mu sigma, and also obtained maximum likelihood estimators for the parameters mu sigma, and then hence both we will talk about the distribution of those statistics, wherein we will encounter the distribution and the multivariate form of the t distribution that is hotlinks t distribution.