Applied Multivariate Analysis

prof. Amit Mitra
prof. Shramishtha Mitra
Department of Mathematics and Statistics

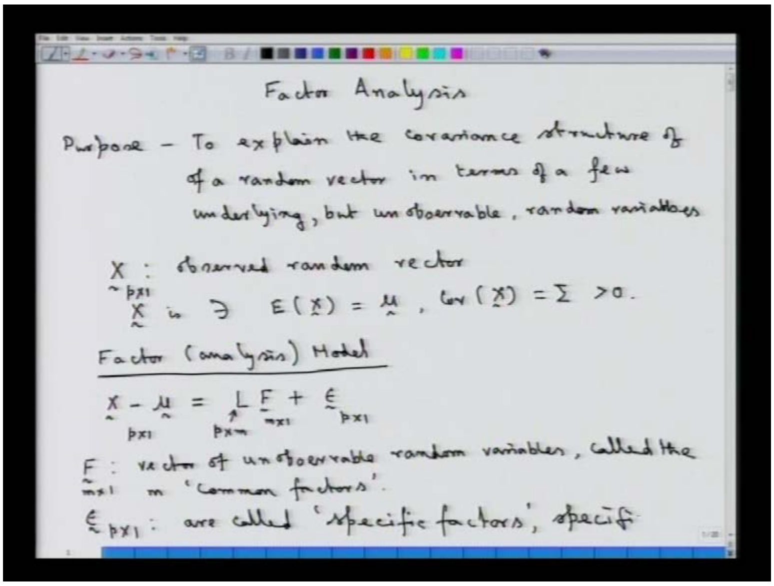
Indian Institute of Technology, Kanpur

Lecture No. # 37

Factor Analysis

In this lecture, we are going to start the concept of factor analysis. Now, the original development of this factor analysis which is one of the important applied multivariate statistical techniques is due to the psychologist, spearman and Thurston. Now, before we actually go into the mathematics of factor analysis - if the factor analysis is an important technique as I said. The basic purpose of factor analysis is that we try to explain the covariance structure of a multidimensional vector, through a through a a number of underlying factors. These underlying factors are what we call unobservable random factors.

(Refer Slide Time: 01:00)



So, let us try to get into this concept of factor analysis. (No audio from 01:03 to 01:10) So, the basic purpose is, what I just now said is that to explain the covariance structure or the variance covariance structure, covariance structure of a random vector of a random vector, in terms of a few underlying unobservable factors. Now, these unobservable factors are random variables. Now, there are points why it is written that in terms of few underlying. Suppose, we have got X - this is the original observed random vector that we have. Now, this X is such that, expectation vector of X is equal to say mu. And the covariance matrix of this X random vector is given by sigma; we assume as before that sigma is positive definite.

Now, this the factor analysis is concerned about this variance covariance structure sigma, and it tries explain this covariance structure, in terms of a few underlying. But unobservable random variables, random variables. Now, if we look back at principle component analysis. In principle component analysis, also we were concerned about the variance covariance matrix there - now in principle component analysis what the objective was that we had tried to do the following, that it was data dimension reduction without loss of information.

So, we can we were concerned about the sigma the variance covariance matrix, and tried to explain the total variability in the data which is measured through the trace of the sigma matrix. And tried to capture that total variation in X, the trace of sigma matrix through a number of linear combinations of the X random variable components. Now, there actually what we did was we had from X, we had made a made an orthogonal transformation to another space, where we were looking at the space of principle components. And there through appropriate number of principle components, we had tried to capture the total variability in X through a fewer number of linear combinations of X, which where orthogonal among themselves

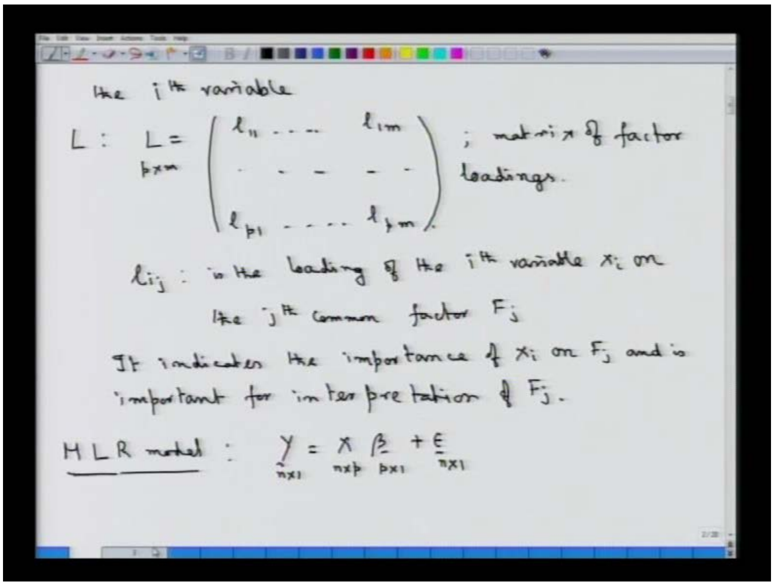
Now here, also we are concerned about the variance covariance structure, the sigma matrix, and we will try to explain this covariance structure sigma. In terms of a few underlying factors, we will call them as common factors. Now, there actually those common factors are something which are unobservable random quantities.

Let us now look at, what what we have as factor model - factor analysis model. The factor analysis model is something like this, that we write X minus mu. So, X is this observed random vector, mu is the mean vector corresponding to that. This is written as an L F plus epsilon going to say what these are. Now, suppose this p p. p is the dimension of this observed random vector. So, this side we have a p cross 1 vector. Now, this L is a matrix of constants, which is of the order of the p by m. This F is the vector which is of the order m by 1, and this epsilon is a random vector which is going to bep by 1.

Now, what are these quantities L F, and epsilon - these are the terms which actually define the factor analysis, the factor model. So, here what we have is this F, which is the vector of our unobservable random vector, unobservable random variables called the m - this is an m by 1 vector. So, m common factors, we will have later on define some actually; we will put some assumptions on this F vector. In order to quantify what sort of random variable that is… Now, the second thing is epsilon - epsilon is a p by 1.

Now, these common factors when we said that F is a vector of unobservable random variables called m common factors. These m common factors are common to each of the original X p dimensional random vector.

(Refer Slide Time: 07:05)



Now, epsilon this is p dimensional vector. This is these are called specific factors, these are called specific factors, specific to the i th variable. And the third quantity as in here in this factor model. So, this F is what we have as the vector of common factors, this epsilon is the vector of specific factors, the thing that is left is L. So, this L is given by this is a p by m matrix of constants. So, this is l 1 1, l 1 2, l 1 m, and this is an l p 1 to up to L p m. So, this L is a matrix of factor loadings. Now these factor loadings has got this interpretation that this l ij is the loading of the i th variable, i th original variable variable

X i on the j th common factor, I am sorry j th common factor which is F j. Now, this has got an important role to play, it indicates this eight of course, is L ij which we have p m of them. It indicates the importance of X i on the j th common factor F j, and thus is important for interpretation of these common factors right.

Now, this is what thus is the factor model - the factor analysis model is so, we write it write the X observed random vector, what we have as the observations this X minus mu is written as L F plus epsilon. Where in this L is a matrix of constant which is, which we are calling as the factor loadings matrix. And we have F the vector of unobservable random variables called the common factors, and epsilon are the vectors. So, we have here, it is better to write it this F m by 1 is this vector F 1, F 2, F m transpose. And this epsilon is the vector - the specific factor vector which is having these p components which are epsilon one, epsilon two up to epsilon p right

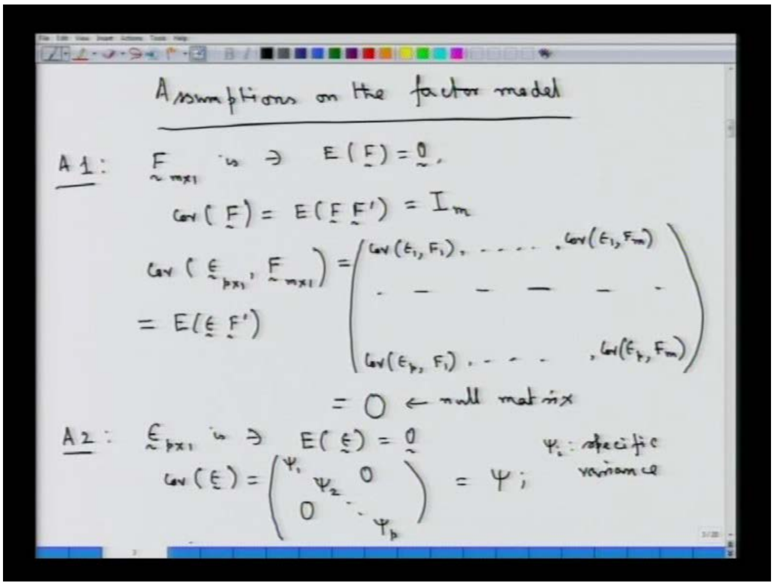
Now, if we look at this factor analysis model this does remind us of a model in statistics which resembles this. If we look at a linear regression model, a multiple linear regression model; a multiple linear regression model looks like the following that, we write Y equal to X beta plus epsilon. Now, in this setup Y is the setup response variables or rather these are observations in the response variable, say this is n by 1 - n observations. And beta is the vector of parameters which are involved in the multiple linear regression model. So, this say is thep by 1, dimensional vector of constants. This X is what is called the design matrix in multiple linear regression, this is of the order of n by p matrix of deterministic constants, and this epsilon is the vector of noise random variables.

So, this is how the multiple linear regression model looks like. If we look back at the factor analysis model, there is the similarity between the two models. Here also if we define this to be equal to some vector Y; then Y is equal to L F plus epsilon. We have in the multiple linear regression model Y equal to X beta plus epsilon - X is the matrix of constants. Here in the factor analysis model L is the matrix of constant.

So, what is the difference between this multiple factor analysis model, and the multiple linear regression model. The difference is that in the multiple linear regression model, we have X to come up with the values actually of independent random variables, which are observable. And these are, this is the vector of constant which are to be determined through the multiple linear regression model. And if we look at this factor analysis model, although it looks like the multiple linear regression model, there is the huge difference between the two models as such. Here we are having F to be the vector of this random variables which are unobservable.

So, while a multiple linear regression model tries to explain this Y response variable through p plus n random vectors, where in epsilon is the vector of noise, and X corresponds to the values of a random variables which are observable. In case of a factor analysis model, the right hand side none of them are observable. These are vector of specific factors or simply the noise. And epsilon this F vectors is the vector of random variables which are vector of unobservable random variables, they cannot be observed right, and L is the matrix of constraints right. So, there is a huge difference as such although the two models look similar.

(Refer Slide Time: 13:03)



Now, let us go into the mathematics of F this factor analysis. In order to do that, we require following assumptions; we will put the following assumptionns, assumptions on the factor model. So, the first assumption says A 1 is the following: Now, this F vector which is as we had taken this is an n by one vector. So, this n by one vector is such that, expectation of this F vector is assumed to be a null vector. We also assume that the covariance matrix of this F vector, which does if expectation of A vector is a null vector. This is equal to expectation of F F transpose that is equal to an identity matrix. Identity matrix of order m, thus we have these unobservable random vectors to satisfy this condition, that is they are not only independent. They are uncorrelated random variables, they have got identical variances. So, they have got variance to be one for all of them.

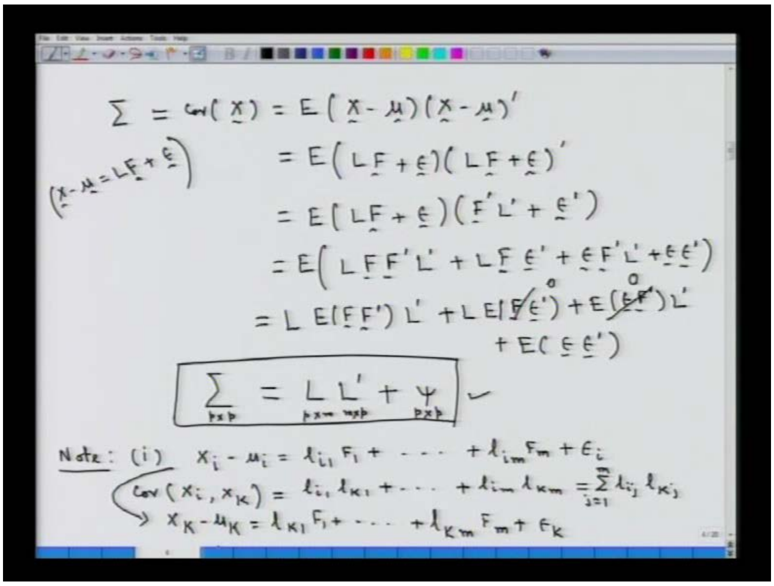
In addition to that, we require the following thing that covariance between epsilon epsilon is our p by 1 vector. So, this p by 1 vector, and this F vector which is m by 1. So, this is going to be a covariance matrix, which is going to be of the order of p m - p crosses m. So, this is covariance between epsilon 1, and F 1 like that. We will have the last element as covariance between epsilon 1, and F m. The other elements can be filled up this is epsilon p - F 1. So, these are scalar quantities scalar – covariance's, this is what we have in here. And this is covariance between epsilon p and F M.

So, this is the variance covariance matrix of this epsilon with F. Now, with the assumption that expectation of F is equal to 0, this covariance is nothing but expectation of epsilon vector, that multiplied by this F vector transpose right. So, we assume that this is a null matrix. So, that is the type of assumption that we usually put in, in for this unobservable random vector or the common factors F, the second assumption is assumption concerning epsilon. So, this epsilon is a p by 1 vector, this is such that we also take this epsilon expectation vector to be a null vector. And the covariance matrix of this epsilon vector to be equal to be a diagonal matrix, having the elements say psi 1, psi 2 has psi p - these are all zero elements.

So, we denote that matrix as without loss of any generality. We denote that equal to psi right. So, we assume that epsilon has got this diagonal structure of its covariance matrix. So, these two are the two basic assumptions. There is a name that is attached to this particular psi i element; these psi i's are called specific variances.

The notation is nothing new actually, because these epsilon 1, epsilon 2, epsilon p. These are specific factors, and these this is the variance covariance matrix attached with this epsilon vector. And hence psi i's are what are called the specific variances right.

(Refer Slide Time: 17:05)



Now, with these assumptions we will move further, and we will see this sigma matrix if one recalls, this is what is the covariance matrix of this X random vector. By the definition of covariance matrix this is nothing but X minus mu into X minus mu transpose right. Now, in the factor analysis model what we have taken is the following, that we have said that this X minus mu is written as L F plus this epsilon vector. So, this is what is the factor analysis model. Now using that, we can write this X minus mu as L F plus epsilon into this L F plus epsilon transpose.

So, this is equal to expectation of L F plus epsilon times transpose of this, which is F transpose L transpose, this plus epsilon transpose. Now, we will take the product and then take the expectation inside. So, taking the product this is equal to L F F transpose L transpose. The second term is L F epsilon transpose, this plus this term into this which is epsilon times F transpose L transpose. And the last term is epsilon epsilon transpose.

Now, we take the expectation inside. Now, remember that this L and L transpose both of these matrices are matrix matrices of constant. And hence this expectation can be taken inside, this is the stochastic part here. So, we take the expectation straight inside. So, what we have is the expectation F F transpose times L transpose, similar argument applied to the second, third, and fourth term leads us to the following observation. That this is expectation of F epsilon transpose plus expectation of epsilon F transpose times L transpose, and the last term is expectation of epsilon epsilon transpose.

So, what is this first term equal to this is expectation of F F transpose. Expectation of F is equal to 0, and hence this is nothing but just the covariance matrix of the F random vector. And what is that, that by the first assumption A 1, we have covariance between covariance of the F vector is expectation of F F transpose, and that we have taken as an identity matrix. So, that this is equal to L L transpose. What happens to a second term, this is expectation of F epsilon transpose. Then thus this is just the covariance matrix of F with epsilon, that is what is the definition. So, we will have this is the covariance between epsilon, and F is written in this form which is also equal to the transpose of that. So, we will have this as a, we had this by the assumption as null matrix.

And hence this is the covariance between F and epsilon, which also is a null matrix. So, this vanishes; this is equal to a null matrices. And so is this, this is also equal to a null matrix. And what we have is expectation of epsilon epsilon transpose, expectation of epsilon being equal to 0. This is just the covariance matrix. So, this is further equal to expectation of epsilon times epsilon transpose. So, this is equal to psi. So, we have in the factor analysis model, this is the important structure what we have for an m factor model to hold, remember our original X was a p dimensional random vector. And if the vector of unobservable random quantities was taken as a p dimensional vector, typically what we would like to have p is of of cardinality to be much lower than that of m, and then actually the actual purpose of the factor analysis will be served.

So, we have this as the important relationship, when we have an p factor model to hold for a for an m I am sorry, this is what the dimension is p. So, this is the structure, what we would require in order to have an m factor model to hold, when we have an original random vector to be a p dimensional. So, this one as we can recall is p by p matrix; this is p by m matrix of factor loading. This is the transpose of that, so that m by p, and this is a p by p diagonal matrix. Now, if you thinks can be noted from such a relationship, and from this particular relationship. We will first look at those before we proceed further, the first thing is what we note from this relationship here, that X minus mu is L F plus epsilon. We can look at the i th component here.

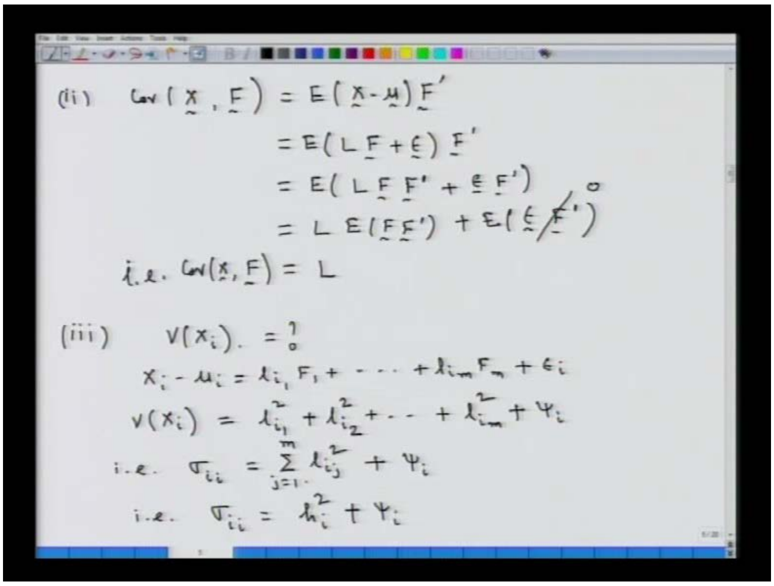
So, we will have X i minus mu i to be written in terms of the i th row here - l i1 F 1 plus l i2 F 2 and so on. The last one will be L im F m ,the m th common factor this plus epsilon i. So, here epsilon i is the i th specific factor - specific to the i th variable here.

Now, once we have this particular relationship here, we can see that what is covariance of X i and X k. Now, covariance between X i and X k of course, is the i k th element of this sigma matrix. Now from this relationship here, what we can see is that if X i minus mu i is equal to this, we will have corresponding to the k th variable - X k minus mu k that would be equal to l k1 times F 1, this plus the k th row. So, this is L k m times F m plus epsilon k.

Now, when we are looking at the covariance between these two quantities. Then since we have F i's, what is the structure the covariance structure of F1, F 2, F n, F m. The covariance structure of F 1, F 2, F m is that, what we have assumed in a one assumption. Where is that? This is covariance covariance matrix of this F is diagonal in nature; with covariance the variance is with of any of the common factors to be just equal to one. So, using that observation what we can say is that this is just going to be the product between these two are quantities, and then taking the expectation. And only this multiplied by the corresponding F 1 term here, this going to have only non-zero contribution - all other contributions are going to be exactly equal to zero.

So, what is this going to look like, this is going to look like that it is L i1, L k1 and so on. The m th term in this sum is going to be L in times L km, this plus what happens here is epsilon I multiplied by epsilon k. And then expectation of that, we have assumed the following structure of epsilon - that epsilon has got this diagonal covariance matrix structure, and hence that also is going to equal to 0. So, we will have covariance between X i, and X k just given by these k terms. Which in terms of summation, you one can write this as L ij times L kj, where j is from one to up to m. So, the covariance between X i and X k which is sigma i i, k element from the sigma matrix is in terms of the factor loadings equal to this term, what we have obtained right.

(Refer Slide Time: 25:04)



Now, another important observations or notes that we can make is what is the covariance between this X vector and the F vector. This X vector is the vector of observable random quantities - observable random variables, and F is the what we have defined as the vector of common factors, which are unobservable. Now, by definition this is going to be expectation of X minus mu, that multiplied by F minus its expectation transpose expectation of F is equal to zero. So, will have this equal to this term. So, using the factor model what we can write is X minus mu equal to L F plus epsilon this multiplied by our F transpose. So, this is going to lead us to expectation of L F F transpose, this plus epsilon times F transpose.

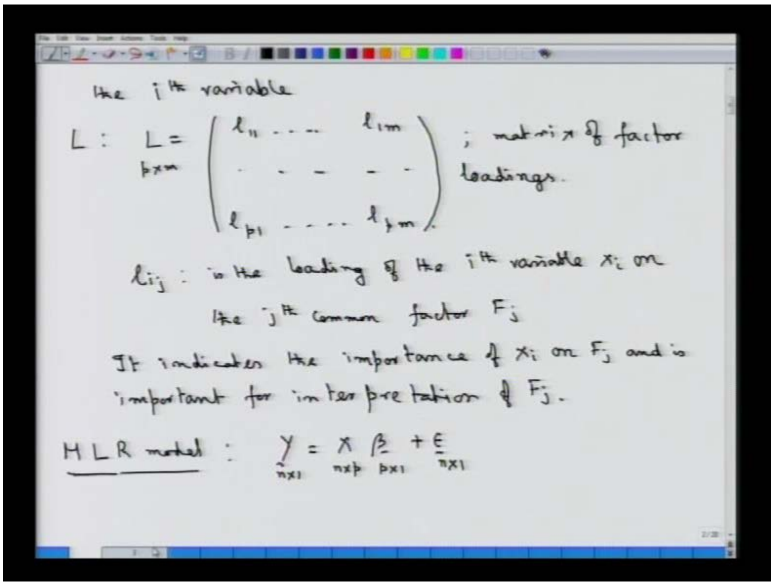
Now, as before this will only be in terms of expectation of F vector multiplied by its transpose plus expectation of epsilon vector multiplied by F transpose. This is the covariance between epsilon, and F which is equal to 0. So, this one is equal to a null matrix, and what we have in here is by the assumption that F has got the covariance structure as identity structure, and hence this is just equal to l. So, the covariance between this X vector, and the F vector of vector of common factors. That is just equal to L right. Now, the third point that we can make note of is that, what is variance of X i? Before we write the expression for variance of X i, we may recall that what this X i minus mu i term was equal to that was equal to our L i1 plus F, I am sorry L i1 into F 1. This plus L im times this m th factor F m this plus epsilon i.

So, when we are looking at variance of X i, it is basically variance of this particular term which is what we have is variance of X i. And then on the right hand side, if we look at variance term by term. Now, each of these F i's are uncorrelated with any other F j for i naught equal to j. And F i and epsilon i that covariance also is equal to zero, because we have taken that way in the assumptions.

So, what we are going to get is that, this is L i1 square. Because variance of F 1 is equal to 1. So, it is just a square here, then L i2 square and so on. The m th term is L im square, this plus variance of epsilon i is the specific variance which is psi i. So, what we have is that this sigma i i which is the i th diagonal element of the sigma matrix, the variance covariance matrix is equal to this is summation L ij, j equal to one to up to m l ij square this plus psi i.

Now, it is written in the following way, because is a special name attached to this one, the first quantity on the right hand side. This is let us write that as h i square this plus psi i. Now, its customary to call this h i square as the i th communality.

(Refer Slide Time: 28:50)



So, this h i square is what we have as the summation L ij square summation from j equal to 1 to m, this is termed as the i th communality. Now what does this i th communality tells us, if we look at this expression here sigma i i. What is sigma i i? Sigma i i is variance of X i , and this i th communality which is h i square is this particular sum.

So, this is the term which is trying to explain the variability in the X i variable, which is the variance of X i sigma i i. So, this is the portion of the variability of X i, that is explained by the m common factors in total. Because this is the contribution in the sigma i i or this is L i1 square is the contribution, in h i square that is coming from the first common factor. L i2 square is the contribution in the i th communality, that is coming from the second factor here.

And similarly this L im square is the contribution of the m th common factor F m in the communality h i square. So, this measures actually the contribution, this is basically the portion of the variance of the i th variable of the variance of the i th variable which is X i, i th variable X i contributed by by the contributed by the m common factors F 1, F 2, F 1, F 2, and F m. So, these are the m common factors, and h i square measures the portion of the variance of the i th variable that is contributed or can be explained by the m common factors, which are F 1, F 2, F m. And the portion of the variance that, the common factors cannot explain goes as the specific variance of the i th specific factor epsilon i, which actually cannot be captured trough the introduced unobservable common factors right.

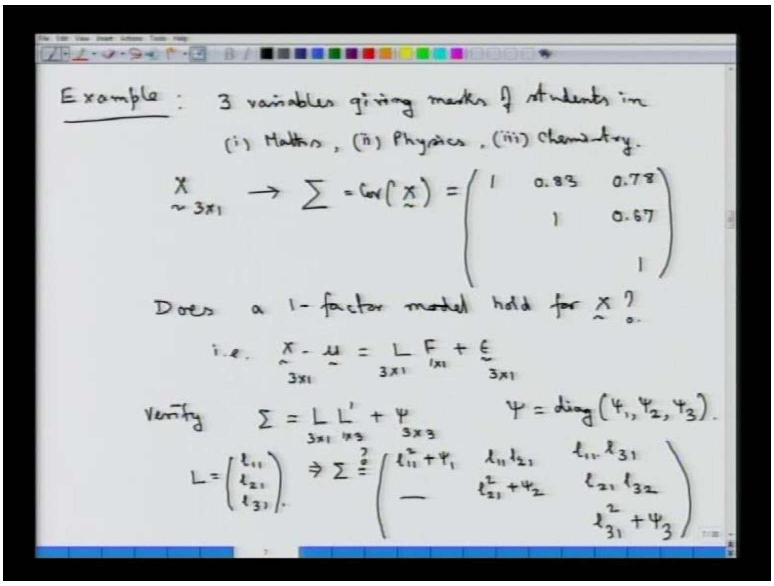
So, this is what is a workable equation actually an important equation that sigma i i equal to h i square plus this psi i factor. Now, we put as the remark the following thing. How we are going to verify this m factor model. Now, we say under the present situation, we say that an m factor model holds for X, if X can be written as X can be written as X minus, that is the factor analysis model L F plus epsilon. With of course, L F and epsilon as formulated earlier. Furthermore an m factor model holds if and only if, model for X holds if and only if sigma can be expressed as sigma equal to L L dash plus psi right.

So, this is how one checks that a given random vector X with a covariance structure equal to sigma. Whether that confirms to an m factor model, where m typically is less than the order of the original random vector that is p, that can be checked using this particular equation. So, this is what it boils down, that a particular random vector X with a covariance structure sigma, we will say that m factor model will hold; if and only if sigma the variance covariance matrix can be expressed as sigma equal to L L dash plus psi.

Wherein as we said, that this L F, and epsilon are as formulated. That would imply that this epsilon is going to be a matrix, which is going to be diagonal with diagonal entries what is the characteristics of that diagonal matrix, with diagonal entries as psi 1, psi 2, psi p. It is going to have those diagonal entries are positive, because this is the variance covariance matrix attached with this psi. Furthermore, if we look at this L matrix the factor loading matrix, we will we had this following relationship that the in note two. We had seen that this covariance matrix between X and F vector, that is equal to L matrix. So, this is that our p by m. So, this will also indicate that what is the covariance between X i and F j, if we look at the i th element here. And the j th element here, this is nothing but L ij - the i, j the element of this L matrix.

Now, this divided by the variance of the standard division of X i is going to lead us to the correlation between X i , and F j. Because the variance of F j terms F 1, F 2, F m they are all equal to 1. And hence this divided by this L ij divided by the standard deviation of X i that is root over sigma i i, that is going to represent the correlation between X i and X j. So, those are the things that would be useful, when we are going for checking whether a an m factor model is going to hold for a particular random vector X or not.

(Refer Slide Time: 35: 31)



Let us look at an example of such a situation. Where we are going to verify, whether a particular dimensional factor model holds for a random vector X. Now, this example deals with three variables - three variables giving marks of students in number one mathematics, number two is physics, and number three is chemistry.

So, we have this X which is a three by one vector corresponding to particular student, which the first dimension gives us the marks of math, second dimension gives us the marks of physics, and third dimension gives us the marks of chemistry say. Now, suppose that this has got a covariance matrix, which is the variance covariance matrix of this X say is given by this. This actually is corresponding to standardized variable, and hence the variances are equal to 1 of each of these variables, this actually corresponds to the correlation matrix. This is 1, 0.83, 0.781, 0.671. We just need to write the upper triangular matrix, because this is the symmetric matrix. So, this corresponds as I said to these three variables standardized, and hence the covariance matrix is what actually we are looking at here is the correlation matrix of the three random variables X 1, X 2, X 3.

Now, we will look at this as the starting covariance matrix, and we will try to answer the following question that does a one factor model hold for x. So, that is the question. We are having three variables, we are trying to verify whether these three variables are such that, that is the sigma matrix of this the three-dimensional variable is such that. We actually can write these three variables in terms of only one common factor. So, we start with this sigma matrix, we are trying to verify whether one factor model hold for X. That is we are trying to write this X minus mu as L F plus epsilon. Now, we are saying that whether one factor model holds, remember that when we had written X in terms of an m factor model. This was the random vector - unobservable random vector of dimension m cross 1.

So, this is just going to be scalar 1 by 1. So, there is a one factor. So, we are trying to say that, now this is three by one. Now, this does L is the factor loading matrix which is three by one, and epsilon is the vector of specific factor that three by one. Now, as we had remarked here, that we are going to say that an m factor model for X holds. If and only if sigma can be expressed as sigma equal to L L dash plus psi with appropriate values of L matrix, and the matrix psi.

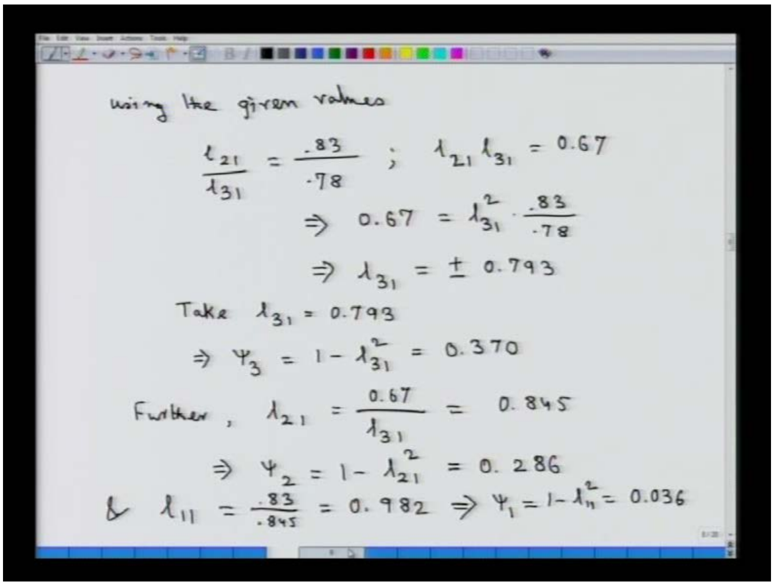
So, we are going to verify this relationship that sigma, whether sigma is conformable to such a representation. Where this is L L dash plus this psi which is matrix I am sorry. So, this is this matrix right. Where this L is three by one, this is one by three, and this is a three by three diagonal matrix right.

So, this L here what we can write is this is L 1 1, l 2 1, l 3 1. So, this relationship here, that sigma equal to L L dash plus psi. This would imply that this sigma matrix, which is what we have no need to write it again. So, sigma matrix is what we have in here, three by three matrix. Whether that sigma matrix now in the right hand side what we are going to do is to multiply L L dash. And then and the psi 1, psi 2, psi 3 which is this psi is a diagonal matrix, with entries as psi 1, psi 2, and psi 3 and thus we are going to write this is L 1 1 square.

So, we are looking at L L transpose. So, this is going to give us L 1 1 square plus the element that is coming from the psi matrix. So, this is psi 1, the second entries going to be L 1 1times l 2 1. The third entries L 1 1 times l 3 1, L 1 1 times l 3 1. The second diagonal entry is l 2 1 square plus this epsilon psi 2, this entry is l 2 1 times l 3 1; and the third diagonal entry is l 3 1 square plus psi 3. Now, this is the symmetric matrix. So, the entry here is just this one, entry here is this one, and this entry is nothing but this entry. So, this is the symmetric matrix. Now, the question is whether we can write sigma with appropriate choices of L 1 1, l 2 1, l 3 1 and psi 1, psi 2, psi 3. Whether the solution of this particular system of equations gives us valid values.

Now, we are going to equate the respective quantity. So, L 1 1 square plus psi 1 is going to be equated with this entry one here. L 1 1 times l 2 1 is going to be equated with this 0 .83, and so on.

(Refer Slide Time: 41:50)



So, the solution is going to be look like the following. So, from the expressions what we get? Using the given values what we get are the following things that, this l 2 1 divided by l 3 1. This is going to be from the sigma matrix this is 0.83 divided by 0.87. Now, this is going to give us or also using this fact that l 2 1 times l 3 1, this is equal to 0.67. What we are doing is basically using all these values. So, what we had first use was L 1 1, l 2 1 that is equal to 0.83, and that divided by L 1 1, l 3 1 which is 0.78.

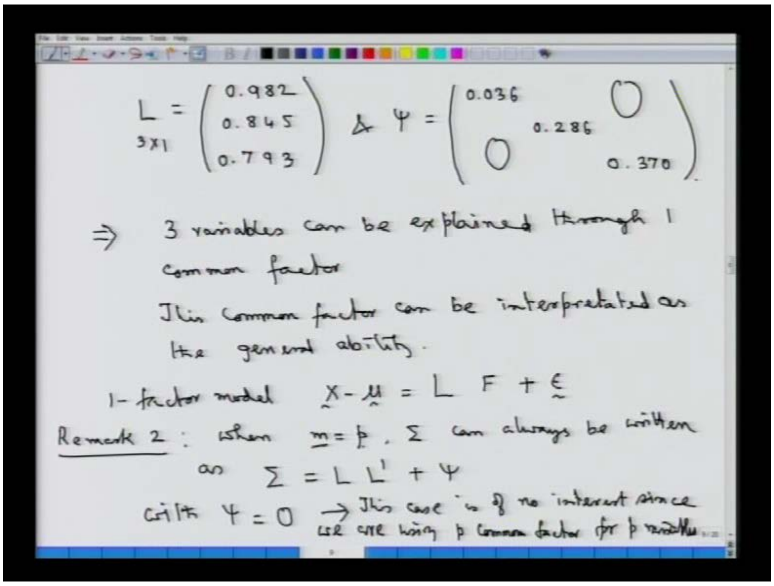
So, using these two values which is 0.78, and 0.83, and 0.78, what we had obtained was this this ratio. Where in this L 1 1 term cancels out. And we have this, and this l 2 1, l 3 1 is coming from this expression, l 2 1 times l 3 1 that is equal to your 0.67. This would imply using these two equations that 0 .67, this is equal to actually this particular entry, that is no l 3 2 element here, this is l 2 1 multiplied by l 3 1. So, this is l 3 1 entry. So, that is this term is now equal to l 3 1 square times 0.83 divided by 0.78. So, this will lead us to the numerical value as l 3 1 square, straight away write this l 3 1 value which is coming out one of the, one of these two values which is plus or minus 0 .793. So, using the two equations we get this solution.

Now, let us take this l 3 1 to be the value with the positive sign, after this example we are going to actually answer the following question that whether, the loading matrix is unique or not. As we will see that, the loading matrix as such when a particular m factor model holds for p is not going to be unique. So, what we are going to take is one of the two values. So, l 3 1 is suppose we take the positive value here as 0.793. Now, from here this will imply that psi 3; now look at this expression here. This term l 3 1 square plus psi 3 that is equal to 1

So, psi three is equal to 1 minus l 3 1 square. So, this is one minus l 3 1 square, and this using this l 3 1 value what we get is this equal to 0.370. So, this square of this quantity subtracted from one gives us this. Now, what is psi 3 - psi 3 was supposed to be the specific variance of the i th specific factor, that was supposed to be positive when it is in this particular case. Further we can solve for the other values, which is l 2 1 can be obtained from this expression, because we now know what is l 3 1. So, we will have this divided by l 3 1 this gives us 0.845, l 2 1 being equal to this this will further imply that psi 2 can be obtained as 1 minus l 2 1 square which turns out to be having the numerical value as 2.286.

And we can also solve for L 1 1, because now we have l 2 1 value also known to us, l 3 1 values also known to us. So, this turns out to be 0.83 this divided by l 2 1 - l 2 1 is 0.845. So, this value turns out to be 0.982 which in turn is going to give us this psi 1 factor which is 1 minus L 1 1 square L 1 1 square. And thus this is going to be equal to 0 .036 right. So, this is entire solution.

(Refer Slide Time: 46:36)



So, after solving this particular equation that sigma is equal to L L dash plus psi or equating these nine values here, what we have obtained let us summarize. So, the solution this is L which is 3 by 1 as we have seen, one common factor. Containing these as the factor loadings which is 0.982, 0.845 this is 0.793 793, and this psi is that matrix of specific variances which are giving us these values. This is first entry this is 0.286, and this is 0.370, rest are null entries. This is what does is the solution.

So, this will imply now since we have in the original setup, this is the variance covariance matrix of standardized variable. And hence the variables, when we talk about the covariance matrix is the correlation matrix. So, these each of these values L ij is l 1 1, l 2 1, l 3 1, each of these now represent the correlation between l 1 1 say is the correlation between X 1 and F 1, l 2 1 is the correlation between X 2. And the only factor what is present there? This is the correlation between the third variable X 3, and the common factor.

And since these are all correlation L ij's are correlations in the present setup with sigma as the correlation matrix. We should have the absolute values of L ij's to be less than one. And what we observe in the solution is that they confirm to that particular restriction, moreover this psi matrix that we have solved has got these diagonal entries to be all positive, and hence this is a valid a one factor model.

So, this implies that the three variables can be explained through one common factor. Now, what is that common factor? If we look at the common factor, then that common factor may be interpreted as the general ability as such. This common factor - one common factor here. This common factor can be interpreted, can be interpreted as the general ability. Now that of course is an unobservable factor as such., and what we have as the one factor model. The resultant one factor model would be an X minus mu that to be equal to this L vector here, that multiplied by this common factor, that multiplied by epsilon.

And hence what if we now want to look at, what is the portion of the total variance of X 1 - what is the variance of the X 1 variable - what portion is explained by the common one single common factor, and what is not explained? The portion which is not explained is this psi 1 factor. Similarly, this single common factor is able to find out 1 minus, this particular quantity of the variance of the second variable, 1 minus this term of the third variable. So, as we see that the first variables variance is largely explained by the first by the only common factor, and it is actually doing this particular explanation of this this order right.

So, this is l 1 1 term square of it, that plus this psi1 is going to give us 1, which is the variance of the first variable as such. Here, looking at these factor loadings one can say that X 1 is closely related most closely related to the factor, because it has got a largest factor loadings which is 0.982. The other scores here, other variables this is what is the corresponding term to the second variable X 2, this is corresponding to third term. And hence, the first variable as such is the variables of the mathematical ability, because we were dealing with this particular a data structure that, the first variable denotes the scores in maths, second physics, and third chemistry.

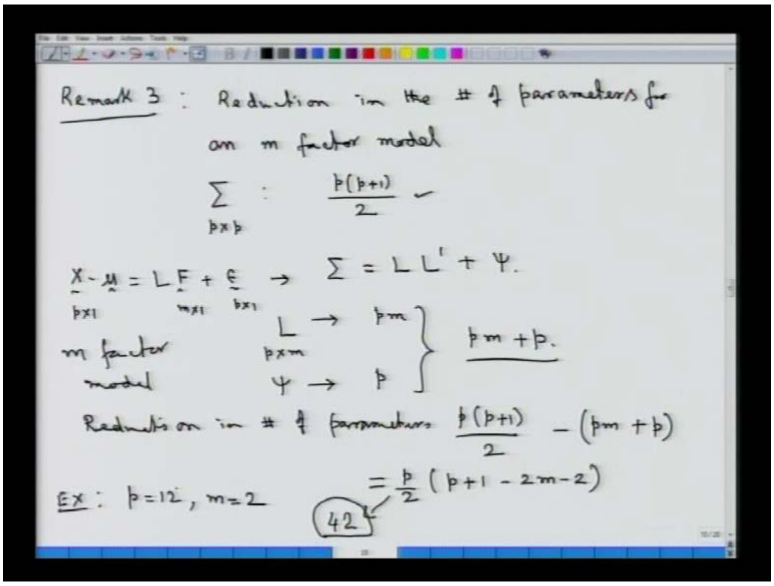
So, what we see here is that the loading of the first variable on the only common factor is the largest. So, it has got, it is mostly it is the it is the closely associated variable to the common factor - the general ability factor the mathematical ability factor.

Now, this is an example just to show whether how to proceed for a given sigma matrix, and to verify whether particular order of a factor model holds for a given sigma matrix or not. Now just look at this particular m factor model, and try to see what type of advantage we are getting, if we are able to express a p variant random vector in terms of an m variety common factor unobservable vectors.

Now, before we do that let me give this particular remark here, we had given some numbers to the remarks, so let me just write this as remark number 1. Then this remark number 2 is that when we take m to be equal to the number of original variables p sigma can always be represented, can always be written as sigma equal to L dash plus psi with psi equal to a null matrix, because sigma is a positive semi definite matrix at least, we can assume that it is positive definite so that there are no no redundancies.

So, this sigma matrix is equal to L L dash that will always be possible; with this psi matrix to be equal to a null matrix; but what we are doing here? We are expressing p variables X 1, X 2, X p in terms of p common factors only. So, there is no reduction as such in this particular situation. So, this case is basically of academic interest; this case is of no interest actually, no interest since, we are using p common factors for p variables only. Now, the actual usefulness of this factor analysis going to come, if we have this m here, the number of common factors is significantly lower than that of the original starting variables.

(Refer Slide Time: 54:15)



And what is that we are going to game, we are going to game the following, let me put it in the remark 3. As in the previous case, we had three variables, three original variables X 1, X 2, X 3 and we had expressed that through one common factor. Now, here what we are concerned about is reduction in the number of parameters; number of parameters for an m factor model. Now parameters in terms of the sigma matrix, because we are trying to explain the variance covariance structure, so we start with this sigma matrix, this is p by p, matrix this is symmetric matrix. So, the number of quantities, so the number of quantities that is there in the distinct element in elements in sigma are p into p plus 1 by 2, which is diagonal plus the upper or the lower entries.

Now, suppose I have expressed this X in terms of an m factor model with X minus mu written as L F plus epsilon, this is say I take this one as before P by 1 and this is say m by 1 common factors. So here, what we are looking at is sigma to be expressed as L dash plus psi. Now, what are the what is the number of parameters that we have in here; in case of this m factor model, this L is what? This L is a matrix of p by m constraints, so this is p by m number of entries, which are there in L; and this psi remember is the diagonal matrix of this p order. So, there are P entries which are there in this.

So, collectively in the m factor model, so this is what m factor model, what we are going to have is this p m plus p. So, the reduction in number of parameters, reduction in number of parameters is this is the original setup, what we have is p into p plus 1, this by

2, this minus this is p m plus p is the number of parameters, which are there in the m factor model. So, simplifying this, what we get is p by 2 into p plus 1 minus 2 m minus 2. Take a particular example, this is the reduction in the number of parameters, when we are able to express say and p variant random vector by an m dimensional model

So, the example here, if I take p equal to 12, m equal to 2, so we have 12 original variables to start with, and suppose we have been able to express those 12 variables in terms of two common factors, then this term here, which gives us the number of reduction in the number of parameters, if you calculate this is this will turn out to be equal to 42, which is huge actually. So, by using a two factor model, two a twelve dimensional original variables, we are able to reduce the number of parameters, which is explaining the variance covariance structure sigma through the m factor model with m equal to 2, we are able to reduce the number of parameters to 42. So, we will continue from this, in the next lecture.