INDIAN INSTITUTE
OF
TECHNOLOGY
KHARAGPUR

NPTEL
National Programme
on
Technology Enhanced Learning

Applied Multivariate Statistical Modeling

Prof. J. Maiti
Department of Industrial Engineering and Management
IIT Kharagpur

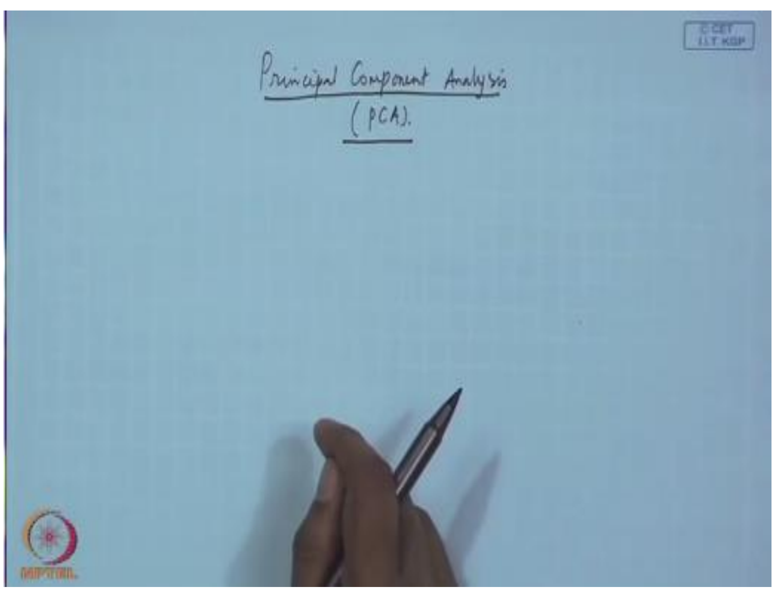
Lecture – 30

Topic

Principal Component Analysis (PCA)

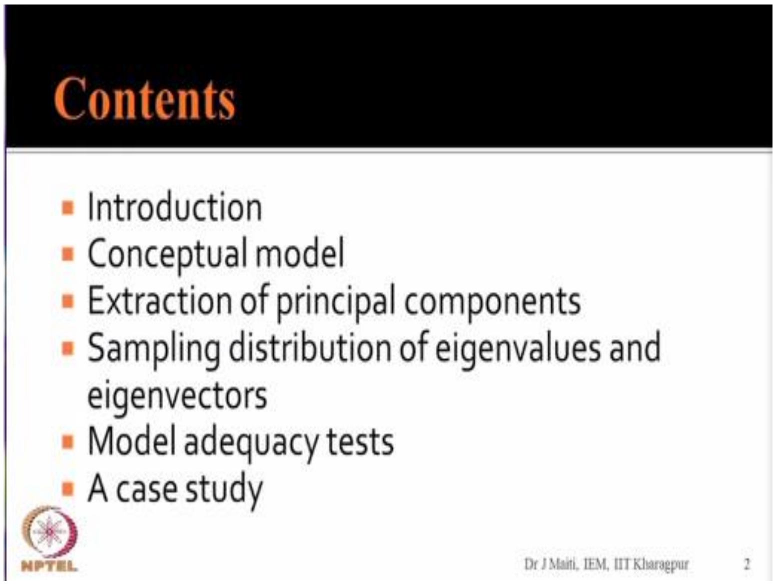
Good morning today we will discuss principal component analysis.

(Refer Slide Time: 00:25)



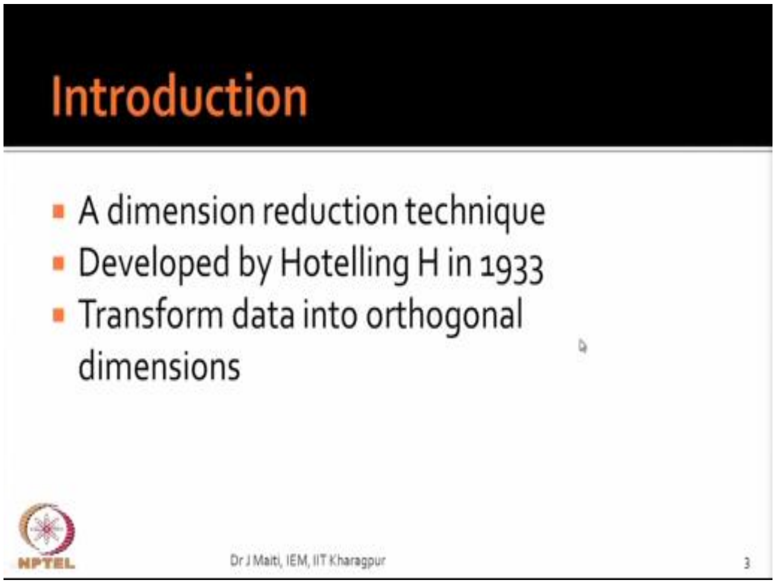
Principal component analysis PCA so let us see the content of today's presentation.

(Refer Slide Time: 00:43)



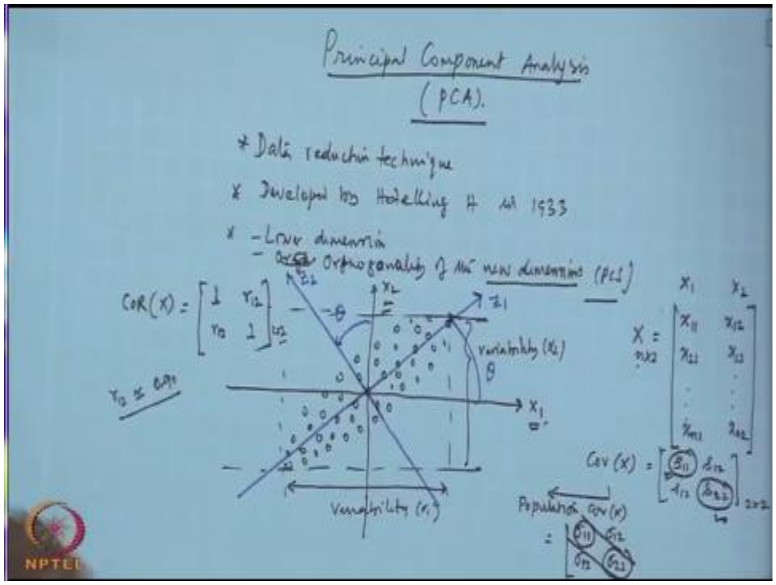
First we will describe the basic concepts prevailing with principal component analysis. Then we will see that how principal component can be extracted from a given data set. Then we will go for sampling distribution of Eigen values and Eigen vectors. You will see that Eigen value, Eigen vector decomposition of the covariance or correlation matrix is the means for extracting principal component. Then followed by model adequacy test, then we will describe one case study. I think it requires around two hour of lecture first hour we will try to complete up to that extracting principal components.

(Refer Slide Time: 01:44)



So, what is principal component analysis so, principal component analysis is a data reduction technique.

(Refer Slide Time: 01:51)



Data reduction technique it is extensively used and developed by developed by Hotelling H in 1933 and it is the data reduction is done with two prospective in mind, one is that lowered dimension, second one is orthogonal orthogonality of the new dimensions. Orthogonality of the new dimensions okay new or transformed dimensions, which are basically PCS principal component. So, let us now go for the very simple two dimensional plot for example, let our dataset is X, which basically composed of n x 2 that means m measurements on 2 variables. So, like this, so this one is variable x_1 and variable x_2 then this one x_{11}, x_{21}.

So, like this xn_1, then x_{12}, x_{22} like this, x_{n2} okay if you plot the data is here suppose this axis x_1, this is x_2. now, if you plot the data let us assume that the data is like this data plot looks like this okay now if you compute the covariance matrix of x. you will be getting as we are talking about data, so we have collected a sample in that sense you have to write this will be s_{11}, s_{12} , s_{12}, s_{22}. So, 2 x 2 matrix so this scattered diagram shows that there is a relationship between x_1 and x_2 so, if you calculate correlation matrix between for x, what will happen you will get certain correlation coefficient.

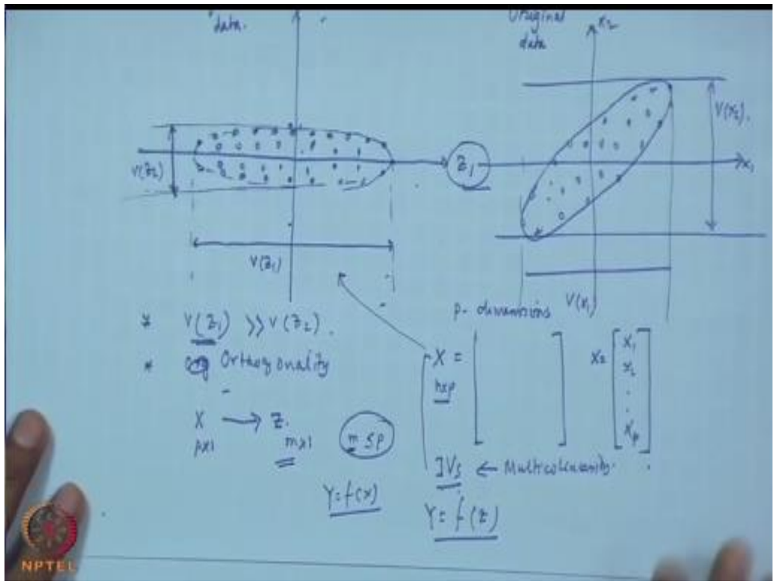
This will be your correlation matrix 2 x 2 you will find out that r_{12} to that will be as you now, that correlation coefficient varies from- 1 to + 1 and here x_2 and x_1 the relationship is positive. So you will get it is a quite large value may be it may be almost = 0.9 here 0.90 getting me. So, this is one aspect that we have taken two dimensional or regional data x_1, x_2 their scattered plot looks like this and it shows there is linear relationship primary linear relationship and if you go for correlation matrix you will be finding out that the relationship is strong enough okay this is first, second issue here if i want to see the variability of $x_1$and variability of x_2 then these are s_{11} and s_{22}.

This is the variance component now if we go for that this is sample, suppose population covariance population co variance, then what will happen this will become σ_{11} σ_{12} σ_{12} σ_{22}. Now, this population variance component for x_1 is this σ_{11} and for x_2 is σ_{22}. If we see the scattered here if we say that n is representative enough for the population large value. If we see the scattered here then you see that the across this x_1, this is the range variability range similarly, if see across x_2, so you will be getting another range that is variability x_2 okay so, you see that definitely the x_2 and x_1 variability are not same, but their substantial values that means if this is my diagram this is my relationship then definitely $s_{11}$and s_{12} s_{11} and s_{22} for $x_1$and x_2 respectively this value is high correct.

Now, let us see that another okay the same diagram, let us see that you are rotating this axis x_1 and x_2 anticlockwise rotating the axis x_1 and x_2 rigid rotation keeping this origin this one visit by angle θ, then what will happen you will get new direction for $x_1$and x_2 correct, if I rotate x_1 / θ then x_2 is also will be rotated by θ. Now, if I say in the new direction this one is z_1 and this is z_2, this one is z_1, then z_2, what can you say about the variability of the dataset along z_1 and z_2 getting me. So, if I draw one more figure along z_1 and z_2 here.

Then suppose, I am writing this one, basically in this, what I am doing now, this z_1 is this along θ, now I am rot I am just making it like this, so that it will be we will be able to see it clearly.

(Refer Slide Time: 10:39)



In the sense correct if you see the scattered plot, now this one is z_1 and this one is other one is z_2 I think it is like this only correct. Now, if we see the variability across z_1, this is variability of z_1 and similarly, if you see this one across z_2 you are getting this variability across z_2. What are you getting here if I plot both the figures now here? So, basically you got something like this one is x_1 and this is x_2 and we have seen that variability of x_1 and variability of x_2. Okay so, this is our original data dimensions and this is as we have rotated by θ degree anticlockwise. We are saying that transformed data getting me transformed data.

Now, it is clear from this two diagram that although here from variability explain variability of x_1 and x_2 point of view. If I see that the both are large and may be x_1 variability little more than x_2, if I go by this $z_1$and z_2, what I am finding out variability across z_1 is much more then variability ax z_2. So, what we have got then variability across z_1 is greater than variability across z_2. Second thing is that this is one second issue is here, that here in $x_1$and x_2 both the variables are co related. That is way you are getting an ellipse by variant l ellipse inclined one that mean this if I see that ellipse here the major and minor axis of the ellipse are not parallel to $x_1$and x_2.

It shows the dependency between $x_1$and x_2. Now here in the transform case, what is happening major axis of the ellipse is across z_1 minor is ax z_2. This shows z_1 and z_2 are independent that is what I said the orthogonal dimensions orthogonal dimension or orthogonality is preserved okay now, can we do this is there any mathematics by, which we can we can transform the correlated structure data structure into uncorrelated data structures well as from correlating that means in the reduced dimension, while reduced dimension if the variability across z_2 is much less compared to z_1, then what will happen we can say that the information content across this z_2 dimension is very less.

We can ignore this information we can just capture the information here in statistical sense information is the variability variance part so, z_1 alone is sufficient to give us the information, what is available in the original data set if this is the case, I will go for only one dimension than this z_1 that this what this the reduction so, principal component analysis we will do this not necessary for necessarily for these two dimension it can be a then for p dimensions by p dimension, what we mean that your x data matrix is n x p that mean the variable you are considering x variable is x_1, x_2, there are p variables p can be quite large, it can may be fifty. So, we are now converting what you are doing this X is converted to z.

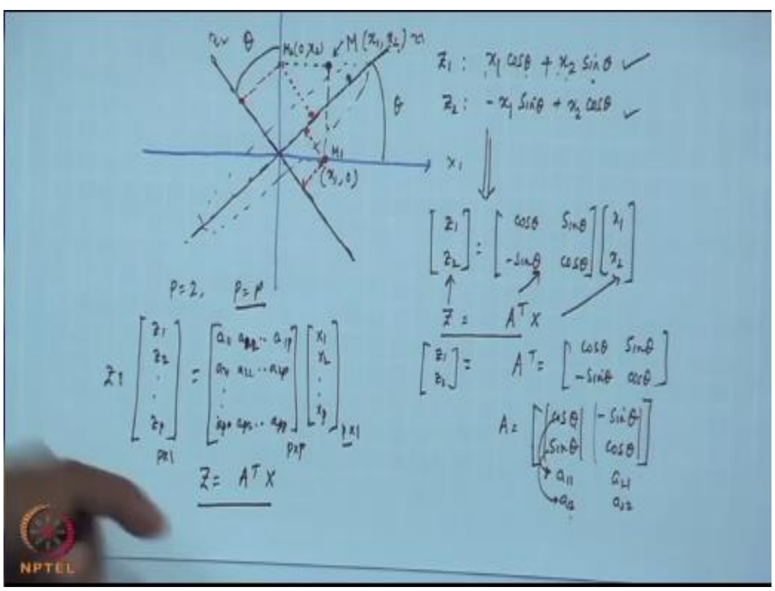
If X is p variable case p x 1 vector z can be p x z also we can extract suppose z is m x 1 vector. So, m can be less the equal to p okay depending on the correlation structure the matrix the matrix will be using the rank of the matrix and all those things, but essentially what do you want basically, we want that m should be less than equal to p. If p is very large m should be as low as possible compared to p. So, that is what is done in principal component analysis getting me. So, principal component analysis is a data reduction techniques it transform original data matrix or data set into some other dimensions of reduced that components and it preserves the orthogonality of the components.

As a result, what will happen what are the advantages you will be getting from here advantages is, suppose we want to do a prediction model using multiple regression and my X variable, these are all IV_S independent variables. If IV_S are correlate then that ultimately leads to multicolinearity problem multicolinearity are there. So, under multicolinearity condition the regression model what you want to fit that y = f(x) linear model. This model will not be a good one, because under multicolinearity it may not be possible to estimate the parameters.

If you use to estimate, what will happen it may be there may be distortion and many things will be there. There are different ways to do for example, is regression can be done in case of multicolinearity problem, but my question is, if I can make them independent, then this IV_S be X can be transformed to this z scale. Like z here these are truly independent so now you can fit a model using f(z) getting me. These are the advantage, now what is the principal basically, how this things that data reduction here or the orthogonal reduction.

That is dimension orthogonality of dimension is maintained and data is reduced to lower dimension what is the method will go by that the two by two matrix case for s.

(Refer Slide Time: 19:17)



So, the same two dimensional case so, if I see this okay so you have seen that we got a ellipse like this in the original dataset $x_1$and x_2 correct now, let us think that there is one point M okay this for example, let this coordinate is x_1 and x_2, x_1 for x_1 variable x_2 for x_2 variable okay now, what you have done when you are converting to z that means z_1 and z_2 so this is z_1 and this one is

z_2.so, I am writing this point little high above for example, let it be here this point okay this is the point.

So, you have rotated x_1 / θ anticlockwise and this also θ I want to know first you say, what is this point this value here, if M is x_1, x_2 then this value is x_{10} for example, let this is M one point on the x_1 axis similarly, if you project this M that projection on this point projection on x_2, what will get you may get a point M_2, what is 0 x_2 this coordinate geometry okay now, what we want we want to find out, what is the projection of this M_1 on z_1 as well as z_2 getting me so, you just let me draw this is axis, now clearly what you want here this is your θ, our aim is we want to find out this is this one x_{10}.

I want to find projection of M_1 on x_2 at z_2 and z_1. So, what you do you will find draw a perpendicular here from this point you want to draw a perpendicular so, from this point draw a perpendicular here similarly, you want also from this point you want to draw a perpendicular line here and perpendicular line here, what you are doing now, that $x_1$component is projected on $z_1$as well as z_2, x_2 component is projected on z_1 and z $_2$.Now, this angle is θ, this is θ, if this one is θ and your this also θ, so this will be θ. So, ultimately, what you will find out that the projection of M_1 on x_1 will be x_1 cosθ.

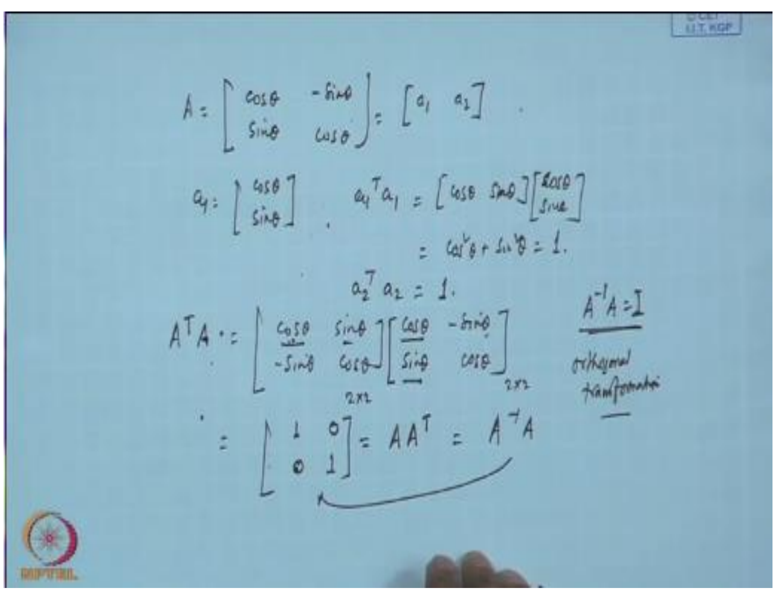
Okay so, if I write z_1 we have there are two projection one is from x_1 cos θ similarly, there will be another projection that is a + x_2 sinθ x_1 cos θ x_2 sin θ okay so, similarly in the z_2 axis, what will happen that - x_1 sin θ + x_2 cos θ, simple trigonometry so, you will be able to find out only just know the θ once you know the θ you know the triangle and project it okay so, then this one if I write in matrix term I can write like this z_1, z_2 = cos θ sin θ - sin θ cos θ[x_1, x_2] we can write like this. Because z_1 = cosθ x_1 x_1 cos θ +x_2 sin θ z_2 = - x_1 sin θ+ x2 cosθ

This equation you are getting okay so, then I can write this one equal to $z = A^T x$ where z is nothing but this one z_1 z_2 and A^T this one. We are saying A^T this is A^T and this is your x and this is your z. So, if i do like this then A^T equal to our cosθ sin θ - sin θ cosθ so, then what is the a matrix transpose of transpose this will be cosθ -sinθ sinθ cosθ this row transformation. Okay so, instead of p = 2 if there is p = p what will happen you will get z_1 z_2 z_p that will be your z and your this equal to A^T. So, a transpose I am now writing that like this our A is a_1 here this is nothing but this column and a_2 is this column.

Okay so, if there are p variable there are two variable, so that why you are getting this one is equivalent to a_{11} this is equivalent to what we are saying a_{11}, x_1 + a_{12}, x_2 so, this is a_{12} and this one is a_{21} and a_{22} you can do this in θ cos θ sinθ we have written here in place of cosθ sin θ I am writing that this is a_{11}, a_{12}, a_{21}, a_{22} like this. If there is p variable case through, you can write down like this a_{11}, a_{21} a this is transpose A^T a_{11} a_{12} like this a_{1p} a $_{21}$ a_{22} like a_{2p}. Similarly, a_{p1} a_{p2} like a_{pp}, this is your p x 1, this one is your p x p and then your x_1, x_2, x_p that means this two dimension this case.

Now, we are making like this and we are writing in terms of s so, this if I write further I can write this one is A^Tx, what we have written earlier also okay then what do you want to see that how this that we say that it will be orthogonal transformation. So, how this orthogonality is maintained here with this two dimension case we want to see.

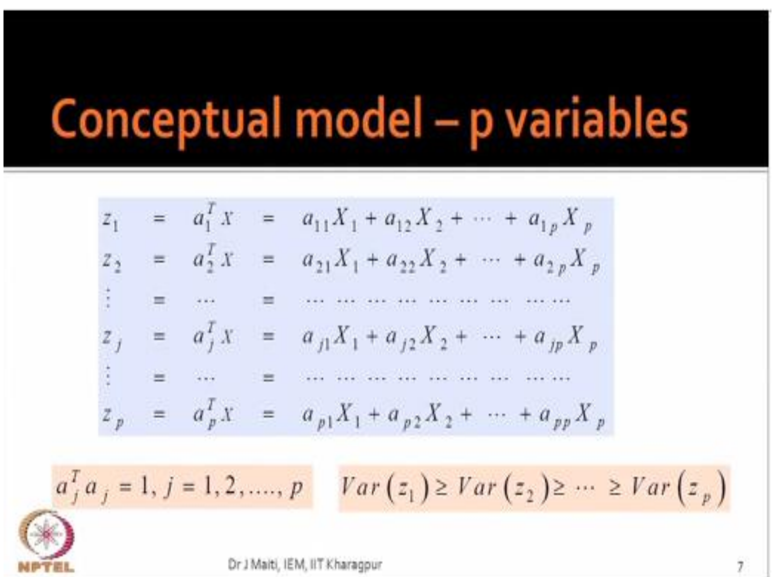
(Refer Slide Time: 28:37)



What we have written for the two dimension case A equal to our cos θ sin θ -sin θ cos θ that we are saying we can write like this a_1 and a_2 getting me. Now, then what is you a_1 a_1 is cos θ sin θ. Now, if I want to know what is a_1^T a_1 what will be the value here cos θ sin θ[cos θ sinθ] will be $\cos^2$ θ + $\sin^2$θ will be one getting me. Similarly, you will be getting a_2^T a_2 = 1 correct now, then if I want to know that what is A^T A that will be cos θ sin θ - sin θ cos θ into a is cosθ sinθ – sinθ cosθ 2 x 2 x 2 cos cos sin sine+ cos ^{2}θ + $\sin^2$ θ that will be 1.Now, then cos – sine sine - cos this will be 0, 1.

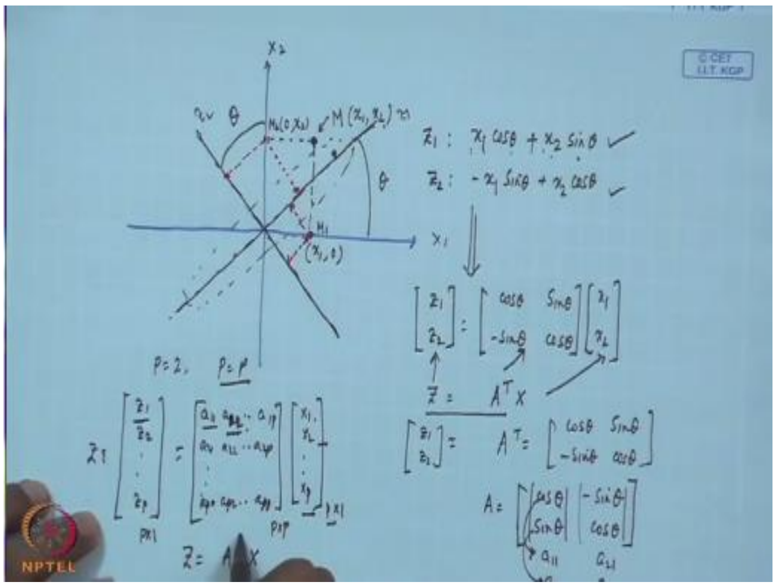
Okay so, this will also be if you do AA^T you will get the same thing correct and see this is symmetric matrix in such a way that if you do like this A^{-1}A is also be getting same thing I mean this is identity matrix. That means what we are getting A^{-1}A is I when A^{-1}A is I this is what is orthogonal matrix that of diagonal elements are 0. So, you are getting orthogonal transformation okay and this transformation matrix, this is orthogonal transformation matrix A. Okay Now, what we will do, I will show you one slide here.

(Refer Slide Time: 32:03)



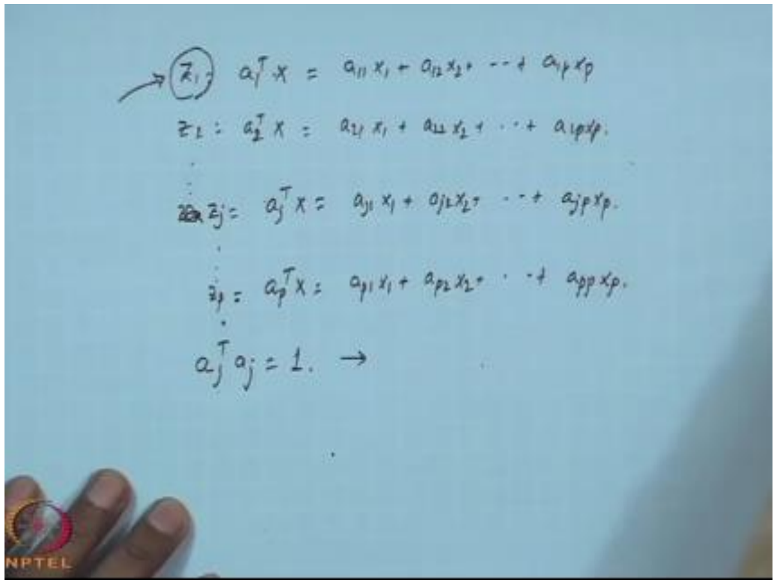
You see that there are p variable for X and you can extract p principal components z_1 z_2 to z_p and their equation will be like this. What we have seen then what I mean to say if you just see this our earlier development that I think, I told you here.

(Refer Slide Time: 32:35)



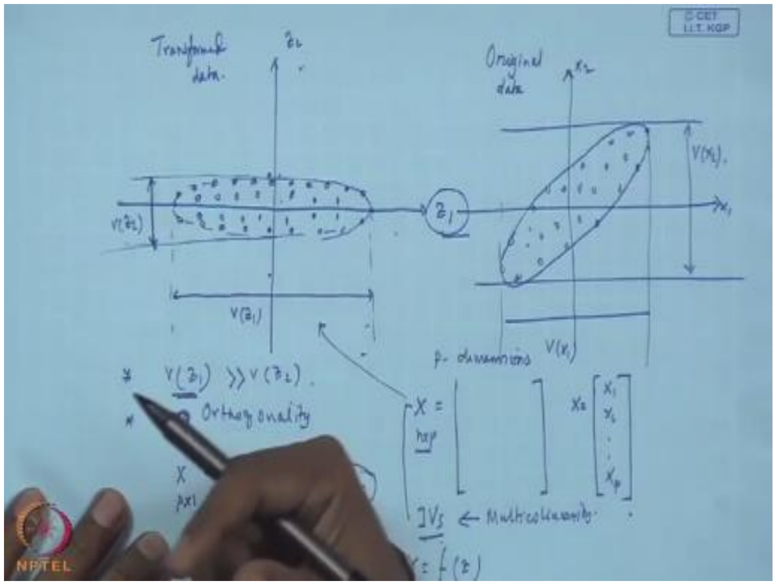
$z_{1= a11}$ x1 + a_{12} x_2 + a_{13} x_3 so, like this you will be getting so what will be getting.

(Refer Slide Time: 32:46)



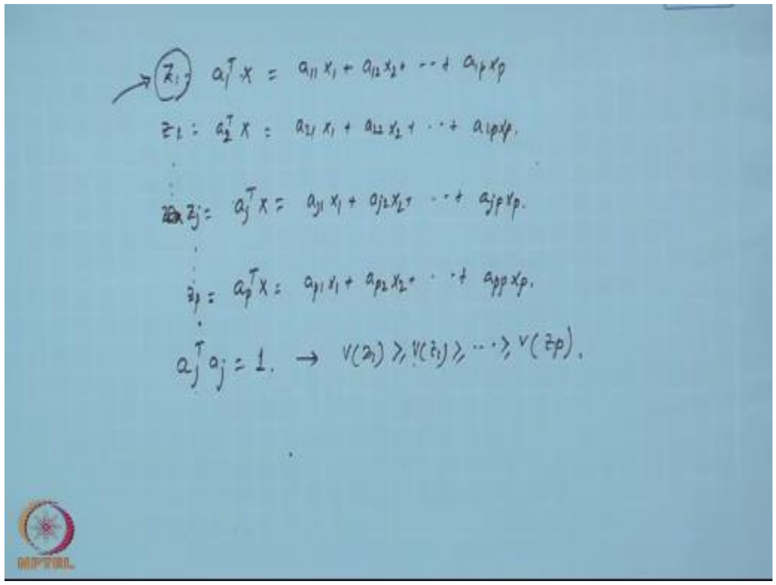
$z_1 = a_{11}$, $x_1 + a_{12}\, x_2 + a_{13}\, x_3 . z_1 = a_1^T x$, which is $a_{11}\, x_1\, a_{12}\, x_2 \ldots a_{1p}\, x_p$ similarly, z2 a2 ^{T} x 21 x_1 a_2 2 x 2 like $a_{2p}\, x_p$ so, if you write down here z_k or z_j you write let it be z_j then $a_j{}^T$ $a_{j1}\, x_1 a_{j2}\, x_2 \ldots a_{jp}\, x_p$. So, following if you go up to the last principal component that may be possible to get that is $ap^T x$, which is $a_{p1}\, x_1 ap_2\, x_2$ like $a_{pp}\, x_p$. okay and we said that we want orthogonal transformation, so when you are reducing the dimension the principal components are, we want this type of a trans aj^T a = 1 that is one thing second one, what do you want we want that when we are extracting the principal component z_1, which is the first principal component the variability in z_1 must be the maximum must be the maximum. What we have discussed earlier in this explanation.

(Refer Slide Time: 34:46)



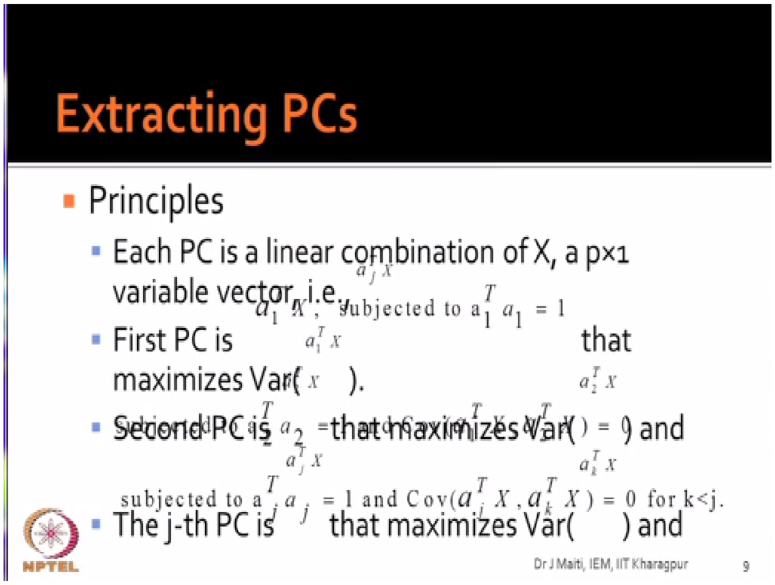
We say we want the first principal component in such a manner that, it will explain the maximum variance possible second principal component will explain the next maximum variance followed similarly, third one like this with this is the case.

(Refer Slide Time: 35:11)

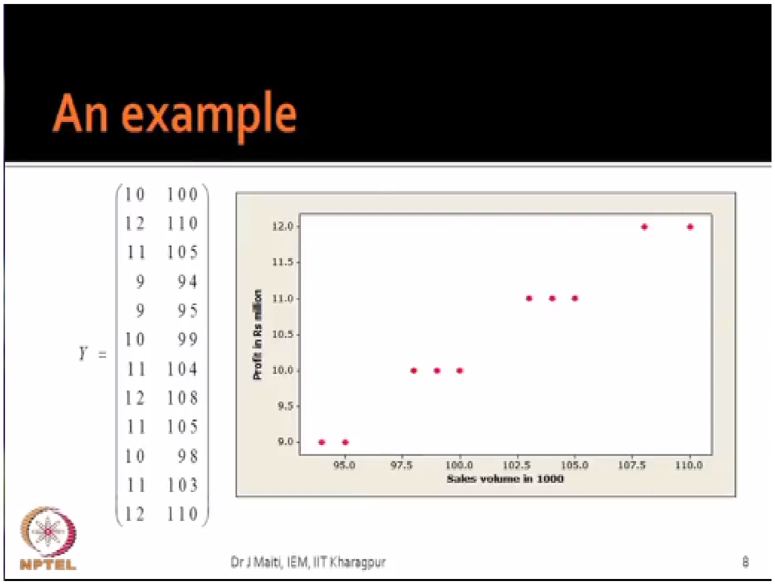


Then I can say here variance of z_1 greater than equal to variance of z_2 greater than equal to variance of z_p.

(Refer Slide Time: 35:29)

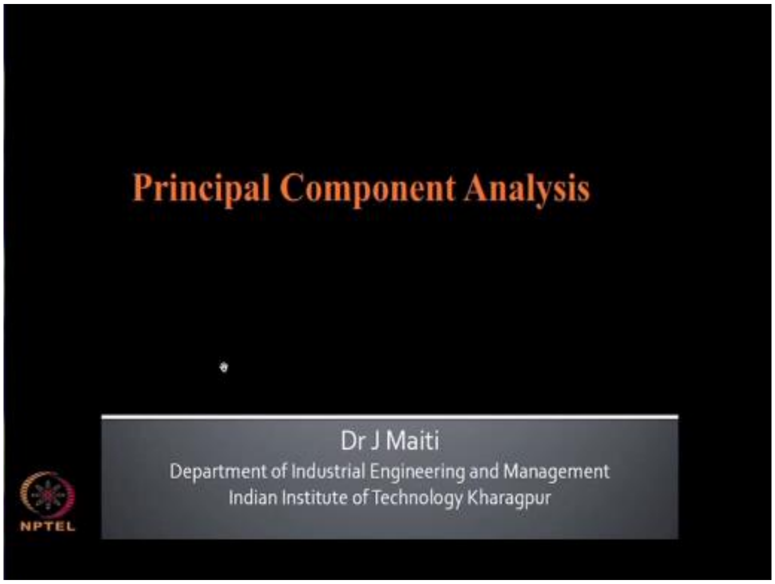


(Refer Slide Time: 35:34)

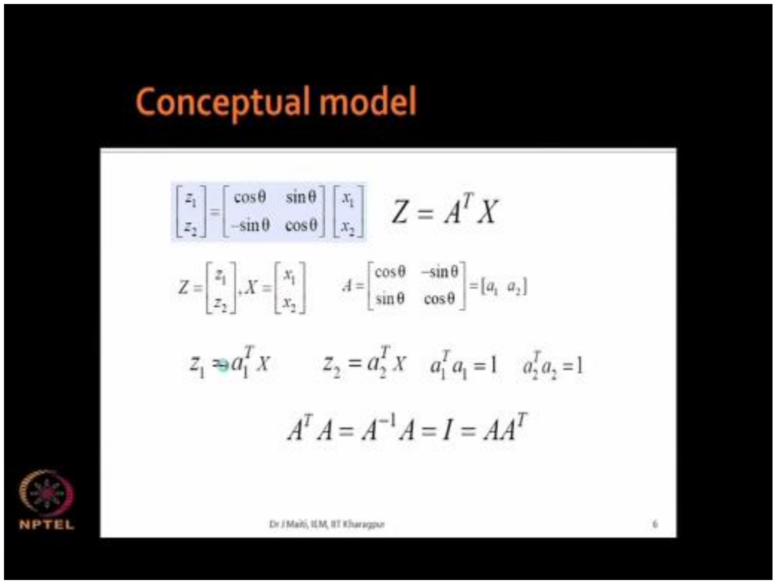


So, what we will do then we will.

(Refer Slide Time: 35:43)

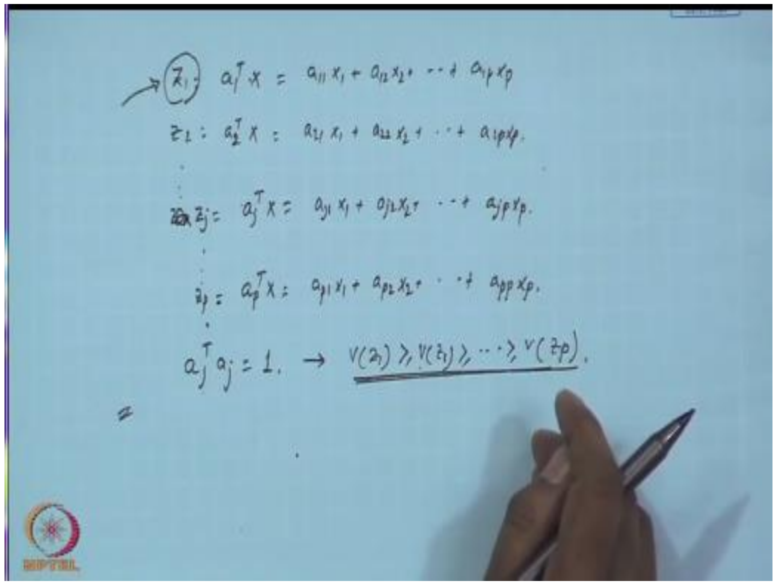


(Refer Slide Time: 35:46)



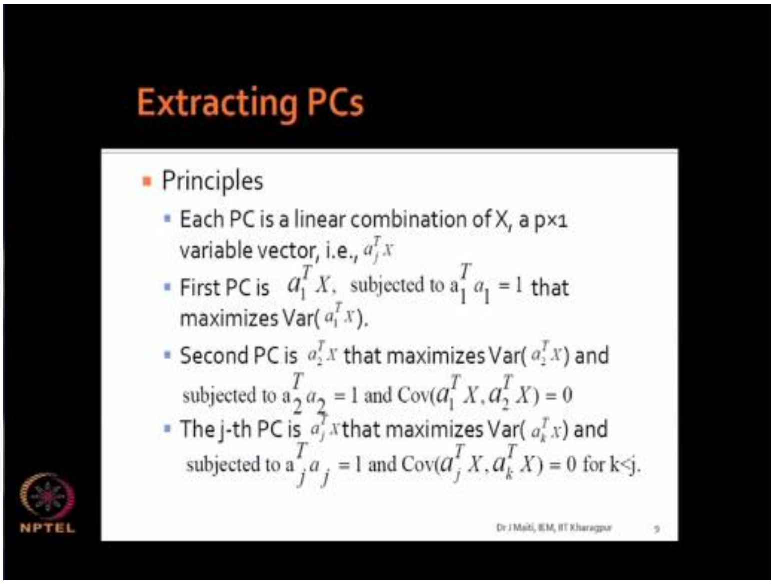
Just one minute.

(Refer Slide Time: 35:53)



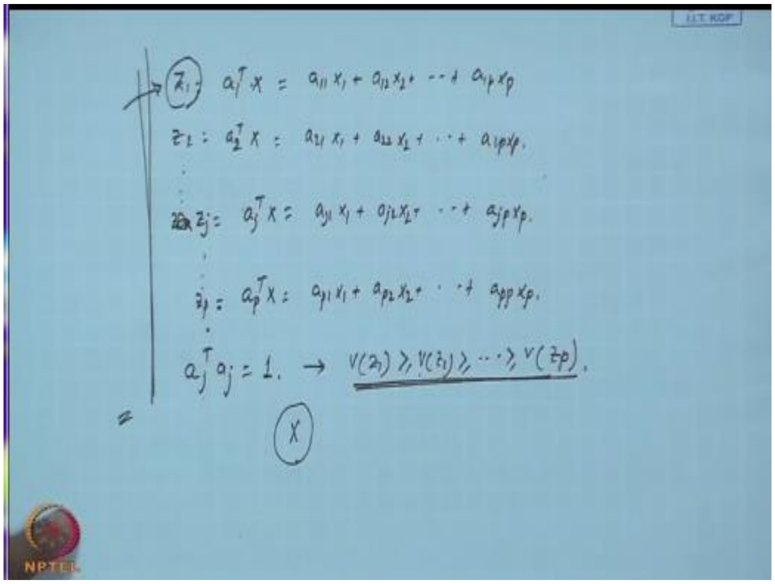
So I have shown you this thing and that, this is possible.

(Refer Slide Time: 35:58)



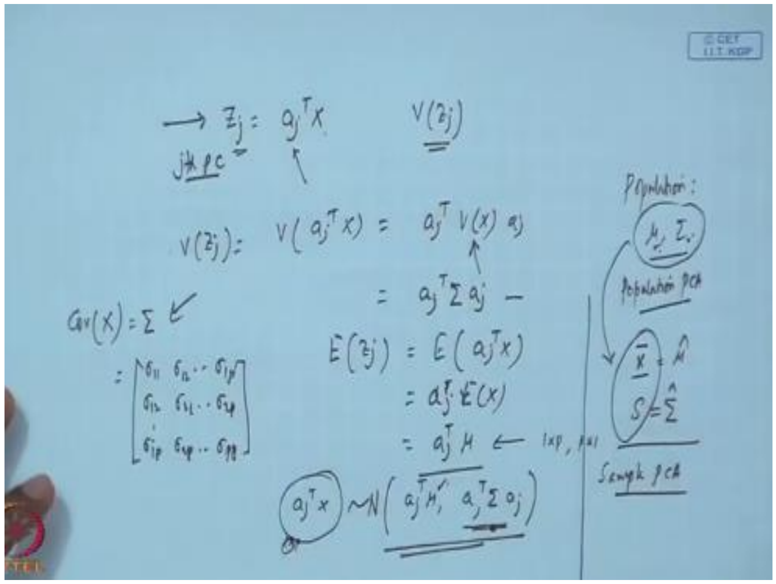
Now, we want to find out.

(Refer Slide Time: 36:01)



What will be the principal components given a data matrix x how to extract the components fine so, here before going to this extraction part I have to tell little bit of that we are saying that in general.

(Refer Slide Time: 36:27)



We are saying z_j is $a_j^T x$ were z_j is the j^{th} principal component this is j^{th} pc and this a_j and x, what are those a_j and a_x you will be x you know what is this a_j, we will be defining late on it is if you see the equation and similar to regression coefficients a_{11} x_1 $+a_{12}$ x_2 + a_1p x_p similar, to regression coefficients, but it is not so, now I want to know I say that variance of z_j you require to compute, because variance of z_1 greater than variance of z $_2$ like this, what will be the variance of z_j.

It will be variance of $a_j^T x$ then it will be a j is the constant here, so a_j^T variance of x a_j. as we know this is a vector basically a_j is a vector later on you will be you know that a_{j1} a_{j1} to a_{jp} that is a vector. Now, what is the variance of x, x is what is the variance of x means, that will be the covariance of x. so, covariance of x will be Σ you have seen earlier indifferent class Σ_{12} Σ_1p Σ_{12} Σ_{22} Σ_{2p} that is Σ_{1p} Σ_2p Σ_{pp}.

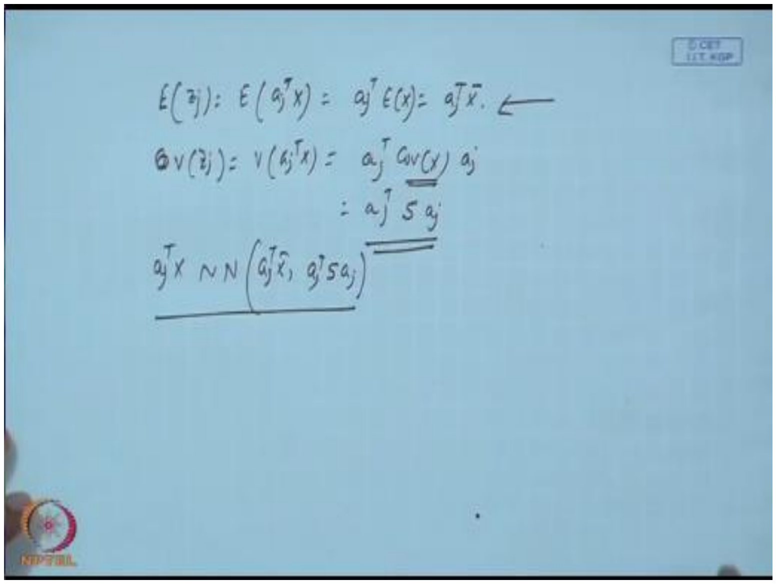
So, that means this will be $a_j^T \Sigma$ a_j okay then what will be the expected value of z_j expected value of z_j will be expected values of $a_j^T X$, which is a_j^T expected value of X, which is a_j^T μ. Because, we said that X is p variable vector and it is it has p mean vector μ_1 to μ_p correct so, that mean essentially that $a_j^T x$ which is something, which is having a_j^T μ $a_j^T \Sigma$ a_j that is the mean and variance part okay so, if there is normally distributed then you will put normal n, so a_j^T is1 x p.

Now, what will be your μ part suppose, this one if you 1 x p into p x 1. So, one into 1 into 1 and this one also will be 1 into p into p 1 into p that is one into1. So, this is a_j^T x this one will give you a univariate case this is linear transformation of variable x all the principal components are linear transformation, that, a_j^T x that is the linear transformation. Now, if you do not get the values of Σ capital Σ as well as the μ, because these are population these are population parameter that is μ and Σ.

If you know μ and Σ and then giving a data matrix you go for this type of extraction of z_j. Then this will be known as population PCA getting me, what I mean to say if you know μ and Σ of the population that x variable case. Then you are doing population principal component analysis, but it is seldom known you will not get this two values. So, what is the then the rescue is you will use $\bar{x}$ as estimate of μ and S as estimate of Σ. So, when we use s actually here, we will be we will be playing with the covariance matrix.

So, if you use S, which is the covariance sample covariance matrix then your principal component analysis will be known as sample principal component analyze. So, we will be writing it as sample PCA and essentially population PCA is not possible. So, we will go for sample PCA, because when we are talking about applied principal component analysis, then we have to rely on the data and will go by sample PCA. So, next we will discuss sample PCA that mean show we will extract all those thing. Okay when we will use sample PCA.

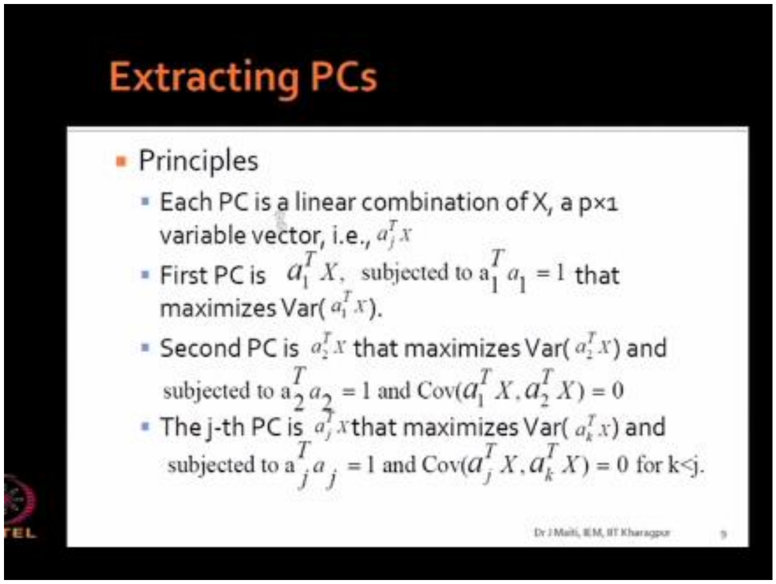
(Refer Slide Time: 42:33)



Then your extracted value of z_j will be expected value a_j^T x. So, this will be a_j^T expected value of x, which will be $a_j^T \bar{x}$ okay from sample point of view similarly, your variance of z_j this will be your variance of a_j^T x, which will be nothing but you are a_j^T covariance of x a_j. This is a_j^T now covariance of S will be replace by S a_j correct so if I say that okay this can be normally distributed then a_jx it is normally distributed with $a_j^T \bar{x}$ a_j^T s a_j okay now, I will explain you when we extract PC.

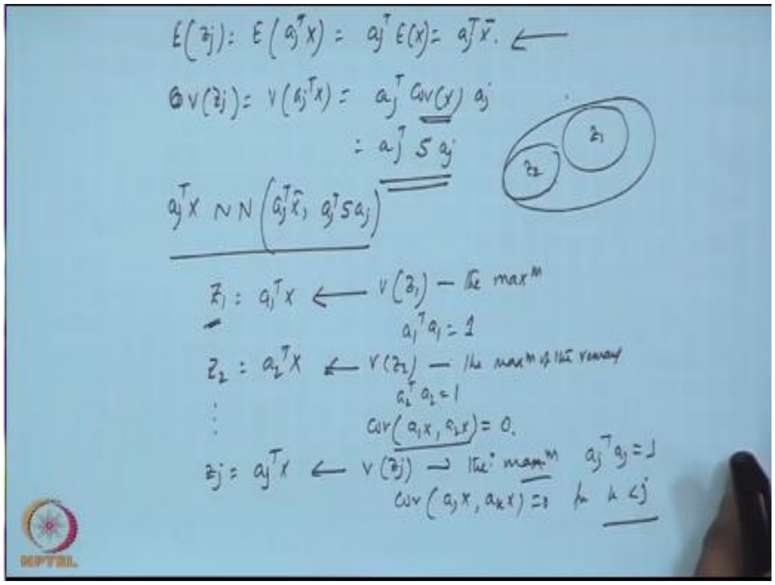
What are the principal we follow getting me, we will we will discuss the principal. Now, let us see the slide.

(Refer Slide Time: 44:08)



Each PC is a linear combination of x_1 linear combination of x ap x 1 variable vector. That is a_j^T x you have seen this one already first PC is a_1^T x subjected to a_1^T a_1 = 1. That also you have describes that maximizes variability of this variance of a_1^T x. Second PC is a_2^T X that maximizes variability of a_2^T X and subjected to a_2^T a_2 = 1. Covariance of a_1^T X a_2^T X = 0, keep in mind this one, what is happening. We are saying first component you are extracting that is our z_1 is the first component.

(Refer Slide Time: 45:14)

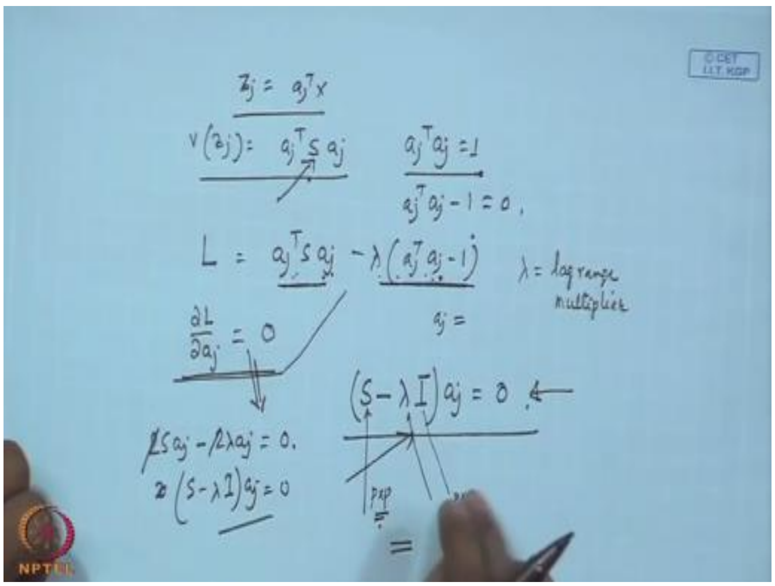


Which is a_1^T X you will extract in such a manner that the variability of z_1 will be the maximum. Also you a_1^T a_1= 1that is the normalization case. So, I have a data set, so and in this suppose this is my total variability portion. So, z_1 will extract as max as possible by z_1.So, let z_1 variability extracted is this much then I am coming to the second PC, which is my a_2^T a x here. What happen the remaining variance I want to extract maximum of the remaining variance so, variance of z_2 that is the maximum of the remaining. So, let I am able to extract this one z_2 see, ultimately if this is my total variability z_1 and z_2 has already may be around 70% is extracted by the two components. There are p components this is also here a_2^T a_2 = 1 and covariance of a_1 x a_2 x this = 0 because, orthogonal component.

Now in this manner if you go for z_j then you will write a_j^Tx. Here, you will definitely maximize the remaining variance already j - 1 components are extracted the remaining maximize this one, the remaining variance. Here, also you will write a_j^T a_j this = 1 and covariance of a_j x a_k x this = 0 for k less than j. Why k less then j what I am saying you're means k up to k can be 1 to j - 1.So, when you are extracting this, please keep in mind this is the principals what I am saying that when you are extracting the first component that basically, takes care of the maximum variance of the data set subject to this normalization.

Then a_2, a_x, z_2 second principal component goes for the second max of variability of the remaining variance and subject to this as well as covariance component between the first two that will be 0 when you go for the third, there also you will you will maximize the variability of that component subject to the remaining variability. And you also make a normalize it that a $a_3^T a_3 = 1$ and covariance between a_3, a_1 as well as a_3, a_2. That x is common that will be 0.So, in the same manner you will extract.

(Refer Slide Time: 48:50)



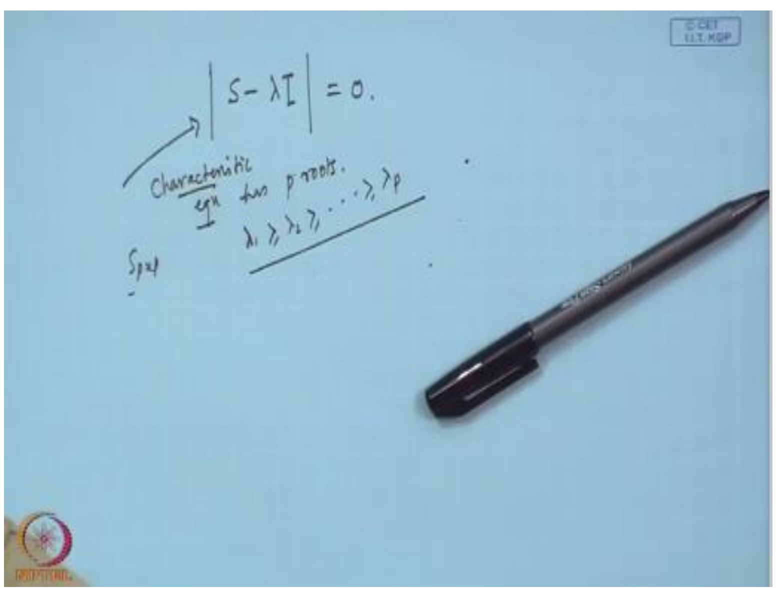
Okay so, then essentially what we are doing now actually, if I go by the general component that is z_j. This is my a_j^T x we are we want to maximize the variability of this, which is our a_j^T S a_j. Now, I am using sample covariance matrix and subject to what we are saying that $a_j^T a_j$ this= 1 correct so, we have two things first one, we want to maximize this second we have this concept. So, this second one I can write like this $a_j^T a_j - 1 = 0$.As we want to maximize this with respect to this condition, so what we can do we can create a function. Suppose, this function is l, which is a_j transpose s a_j- λ into $a_j^T a_j$- 1.

Correct so, we are using a function this is the lλ is the Lagrange multiplier λ is Langrage multiplier. So, you know that this is the general way of maximization process using langrage multiplier. So, what you require to do now you require to find out the a_j value in such a manner that this function will be maximized. So that means what I want δ a / δ a_j, that is what I want getting me, so this you want to put it 0 provided δ2 l / δ a_j that will be that matrix. You will also get this should be that maximization that positive definite and negative definite case is there that condition must satisfy.

So, this ultimately results into this equation it will be like this S - λ I into a_j = 0. See this is if you take derivative with respect to a_j it will be 2 Sa_j if you take derivate with respect to a_j it will be 2 λ a j this derivative. So, ultimately it will be 2 Sa_j- 2 λ a_j that= 0. So, 2 will be cancelled out that0 is there, so S - λ, but s is a matrix λ is scalar this into a_j= 0 so, that is the equation okay so now S is your p x p matrix. Because, p variables are there λ is a scalar it all depends on that what will be there number of λ, but I definitely p x p matrix.

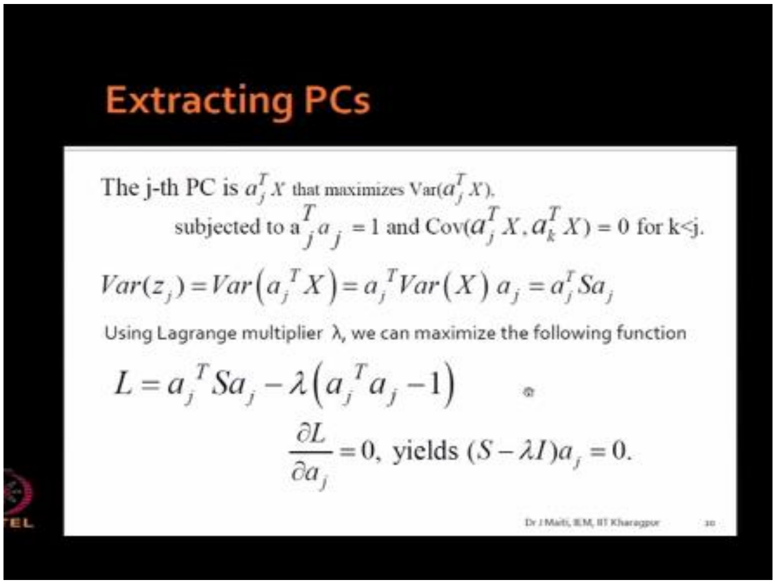
Okay then this equation is a famous equation in matrix algebra, anyways you will find out that set of linear equations case suppose and more famous one is this one.

(Refer Slide Time: 53:04)



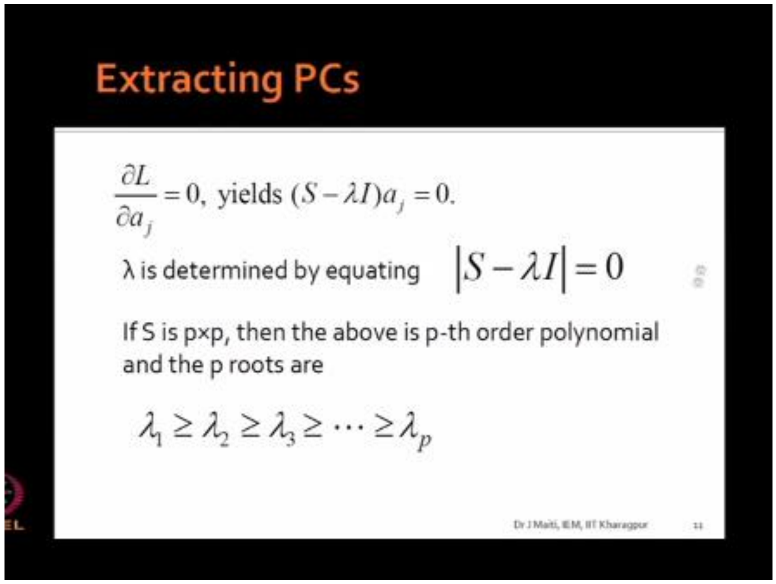
Suppose I want to know the λ values then what you require to do you have to find out the characteristics equation this = 0.So, determinant of this = S – λ I = 0 determinant of this =0. This is a characteristics equation getting me, so if S is p x p then this equation has p roots that means you will be getting λ_1 with the condition greater then = λ_2, like this λ_p.

(Refer Slide Time: 54:07)



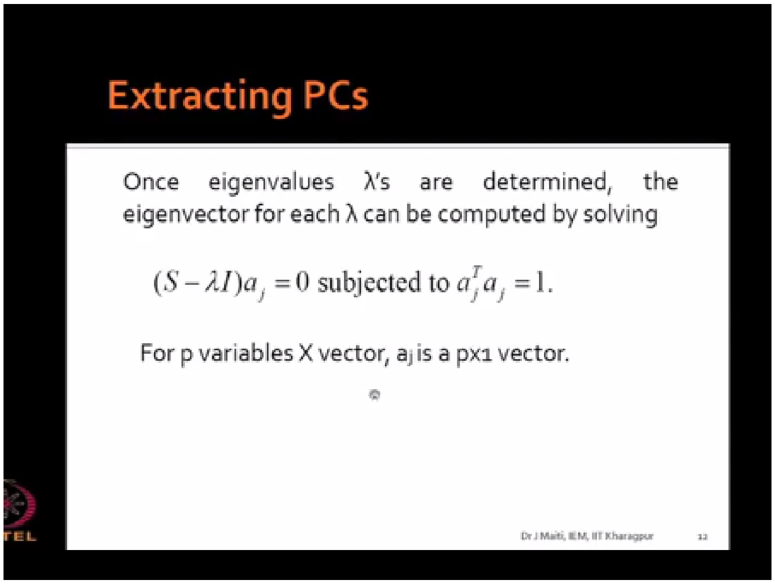
Okay so, let us see now some more slide here.

(Refer Slide Time: 54:11)



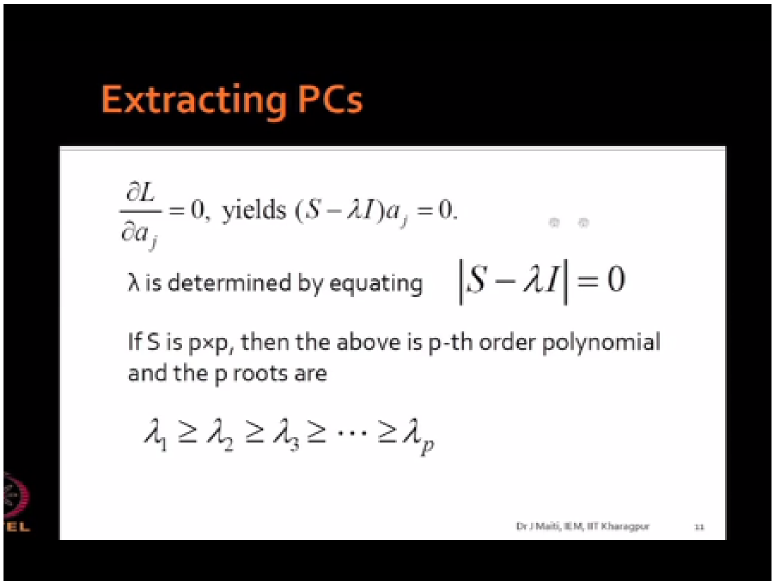
So I said that this will be the that p the order polynomial is there p root you are getting.

(Refer Slide Time: 54:17)



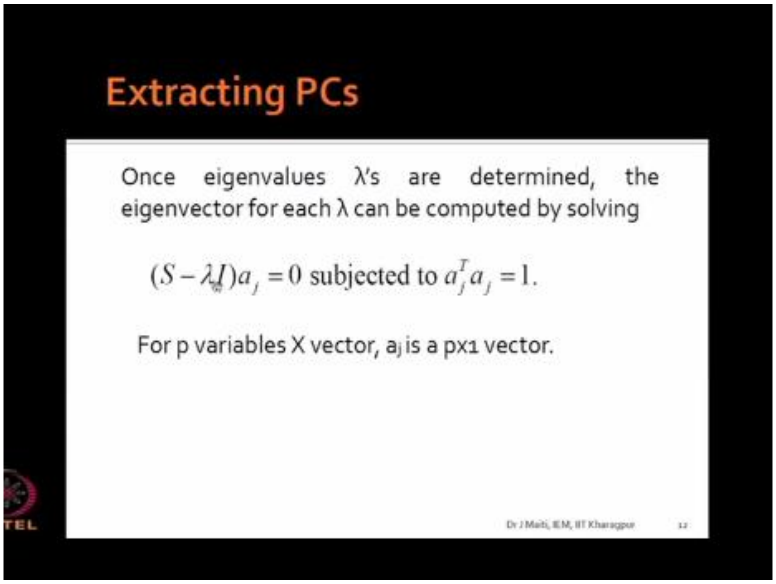
You have to this is the if you know λ.

(Refer Slide Time: 54:25)



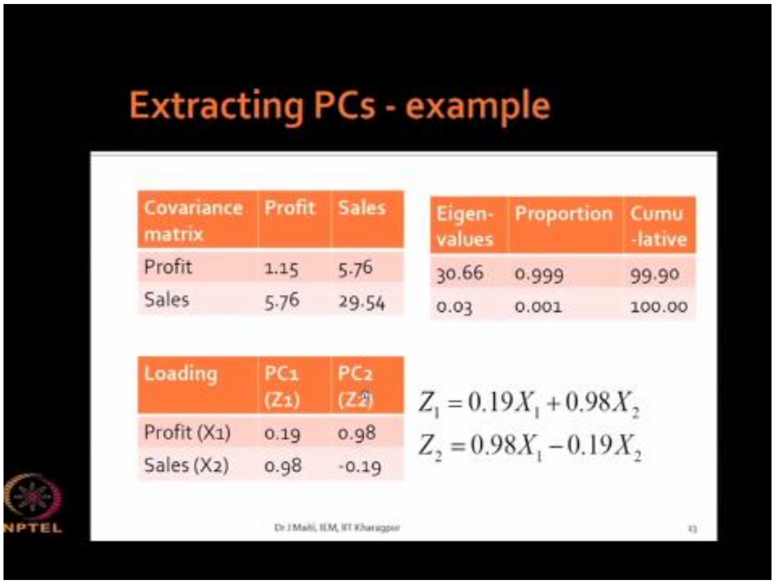
Basically from this equation you will be able to find out λ values. Now, take 1 λ put into this equation.

(Refer Slide Time: 54:30)



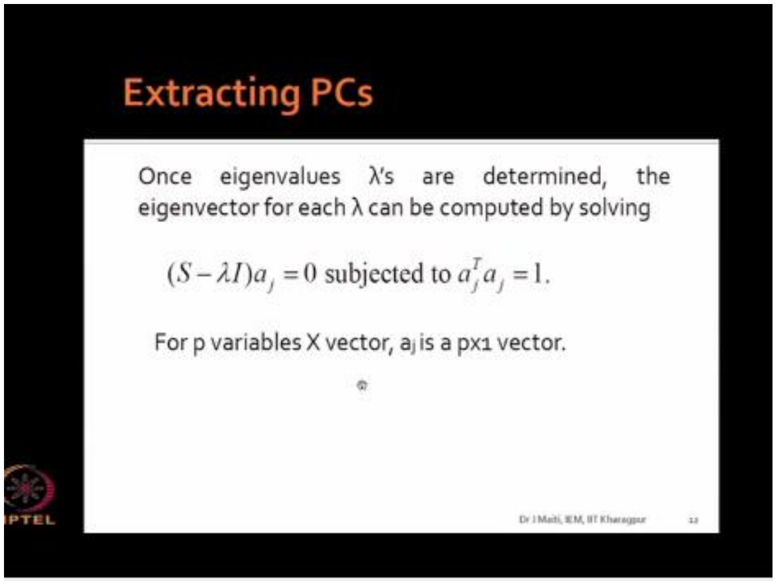
And then you find out the a values okay.

(Refer Slide Time: 54:37)

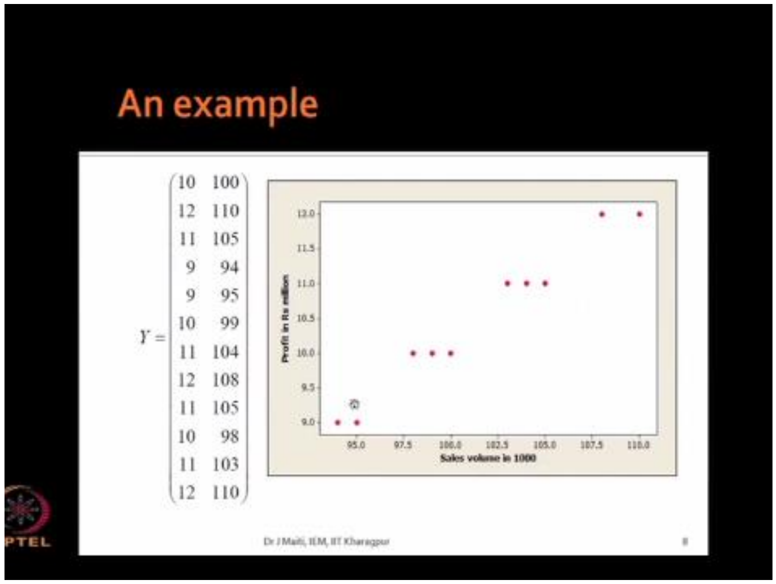


I will show you one data set.

(Refer Slide Time: 54:38)

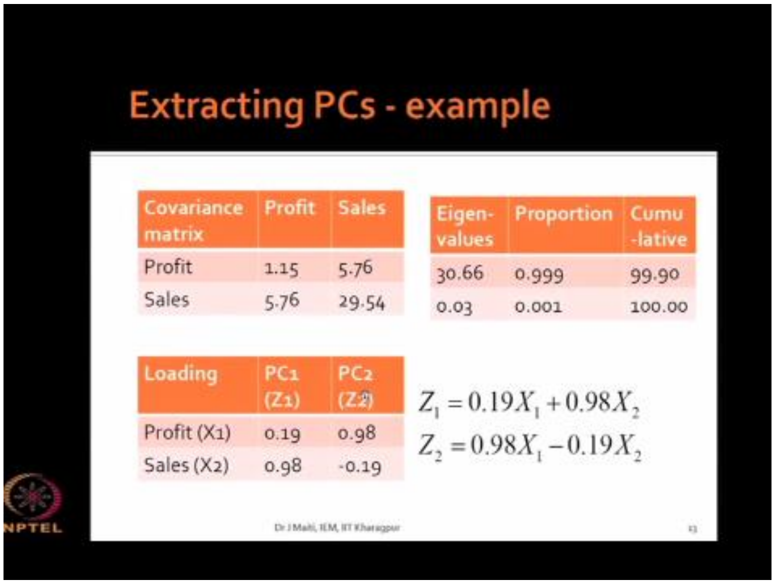


(Refer Slide Time: 54:41)



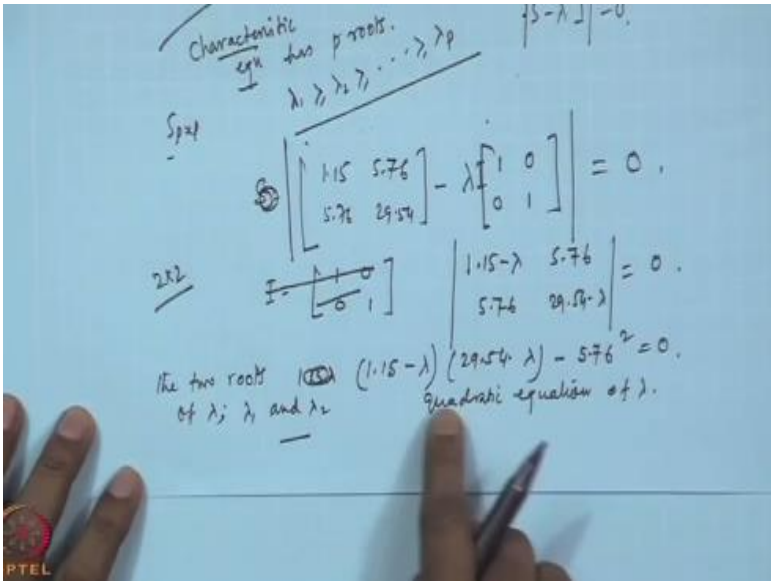
Now, this is the dataset what is the profit and loss? For example we have started with and if you of you plot the scattered, what you are getting you are getting that it is almost linear relationship. So, that means one here data reduction is very much possible, so we tried with two variable case and what we found out that these are the things.

(Refer Slide Time: 55:09)



So, first is the covariance matrix. Covariance matrix is S.

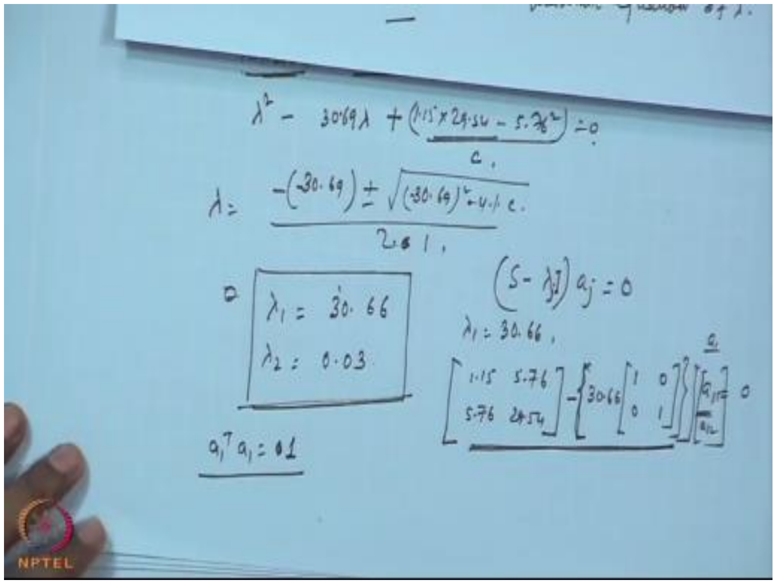
(Refer Slide Time: 55:13)



So with this example S is 1.15, 5.76 then your 5.76, 29.54.Now, what we want we want S - λ I determinant = 0. So your I will be 2 by 1001 I is identity matrix, so if I do like this - λ I. So, that means instead of λ I am writing 1001 okay this total determinant if I write and if I put into 0, then what is the resultant. This one 1.15 - λ5.76 then 5.76, 29.54 - λ this = 0, determinant of this = 0.So, ultimately you will be finding out 1.15- λ into 29.54 - λ -5.76^2 = 0. It will be it is a quadratic equation of λ because our S is 2 x 2.

So, there will be 2 roots, so λ_1 and λ_2, so that roots will be the two roots of two roots of λ. Let it be λ_1 and λ_2 now using this equation you can find out.

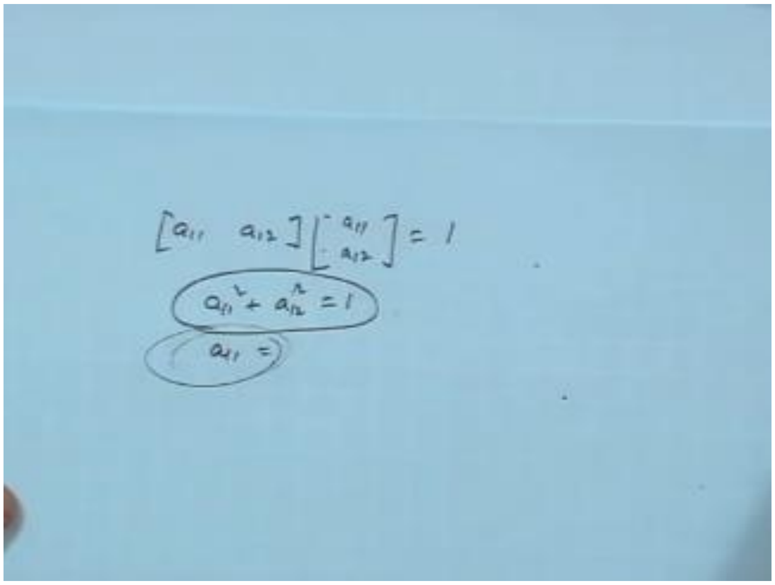
(Refer Slide Time: 57:06)



So, what I means to say the resultant equation is 1.15 into 29.554 - 1.15 λ -29.54 λ + λ^2 - 5.76^2 = 0. So, λ^2 -you add this 30 I think 30.this λ, this one then this two is to be added that you will be getting some value. This is the sum positive value you will be getting okay if I consider this as c then you λ is - b -b is 30.69 - b + - $\sqrt{b^2}$, b^2- 4ac, a is 1 c / 2 into a is 1 this will give you the Eigen values two values λ_1 will be your 30.66 and λ 2 will be 0.03 correct so, you are getting your λ_1 and λ_2, then what you require to find out you require to find out the a values.

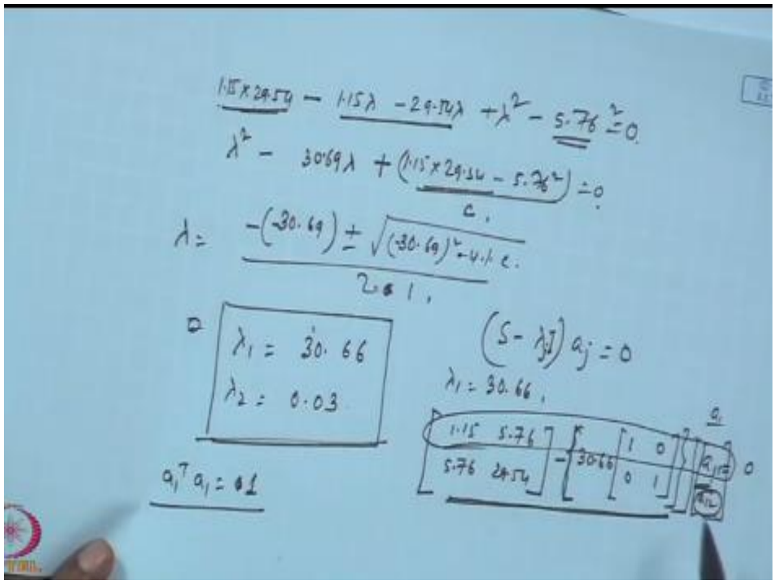
So, what you will do now, we know S- λ_i a_j that will be = 0, so here will be using λ_j. Now, so as λ_1 = 30.66, so you put S is 1.15, 5.76 then 5.76, 29.54 this - λ is your 30.66 and into this matrix 1001. This matrix this into a_1 = 0, now you are a 1 will be what a_{11}, a_{12}, because there are two variable. So, I can write a_{11}, a_{12}, a_{11}, a_{12} that will be 0 this value will be 0, but what will happen ultimately, you will be getting two equation and you will not get unique solution. You will get infinite solution under these conditions because this is a equation is 0 and that means there will be infinite number of this a_1. So, in order to restrict that what we will do a_1^T a_1=1 so if a_1^T a_1= 1.

(Refer Slide Time: 01:01:00)



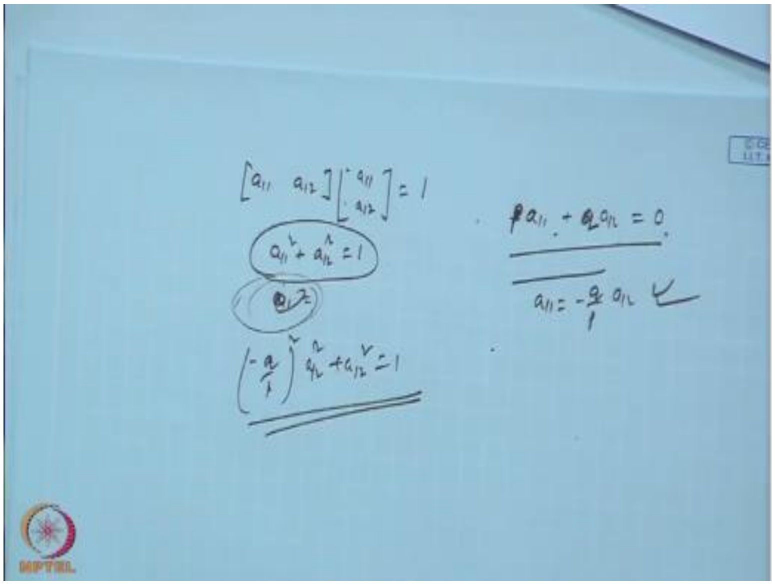
Then what will happen a_{11}, a_{12}, a_{11}, a_{12}that is one. So, $a_{11}^2 + a_{12}^2 = 1$. So, whatever equation you are getting here that earlier equation you put a_{11}= from this equation. What is the value you are getting you will be getting positive and negative values. So, you put this one once you put into this equation.

(Refer Slide Time: 01:01:36)



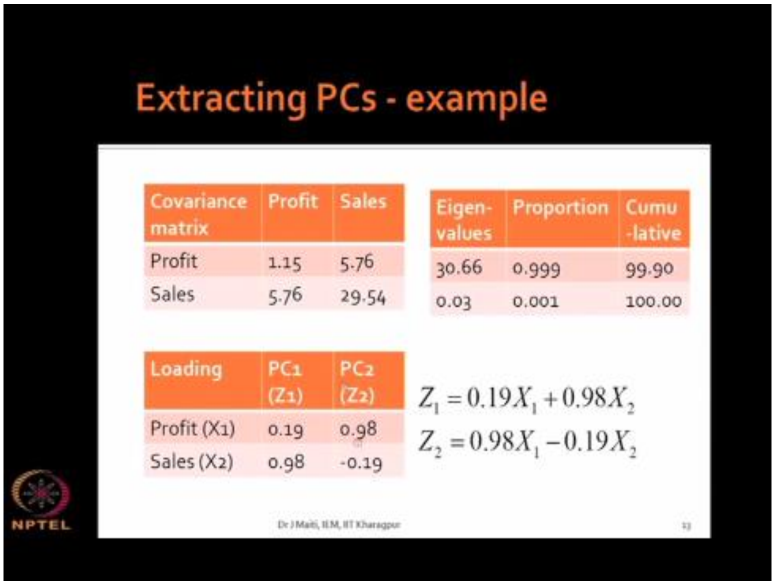
Or a_{11}, now a_{12} will be replaced by $a_{11}$from this first equation a_1 is already there. Our a 1 relation you will be getting from here from this you will be getting, because this is two by two this into this + this into this. So, it is a two variable equation will be like this.

(Refer Slide Time: 01:10:46)



Some value into a_{11}+ some value into a_{12} that will be 0 like this. So, similarly now you are getting one relation from here suppose this is your p and this is your q. Then you are getting a_{11}= q / p - a_{12} Here, you put instead of a_{11}you put - q / p^2 a_1 to + a_{12}^2 s + a_{12}^2 + a_{12}^2= 1. You will be getting a_{12} and putting here you will be getting a_{11}.Similarly, $a_{21}$and a_{22} you will be getting. Okay so, the final one this is the that why this.

(Refer Slide Time: 01:02:51)



This is your a_{11}, a_{12}this is a_{21}, a a_{22} and you are getting third principal component z_1 and z_2, this is the manner of extraction of principal component.

NPTEL Video Recording Team

NPTEL Web Editing Team

Technical Superintendents

Computer Technicians

A IIT Kharagpur Production

www.nptel.iitm.ac.in