Advanced Matrix Theory and Linear Algebra for Engineers

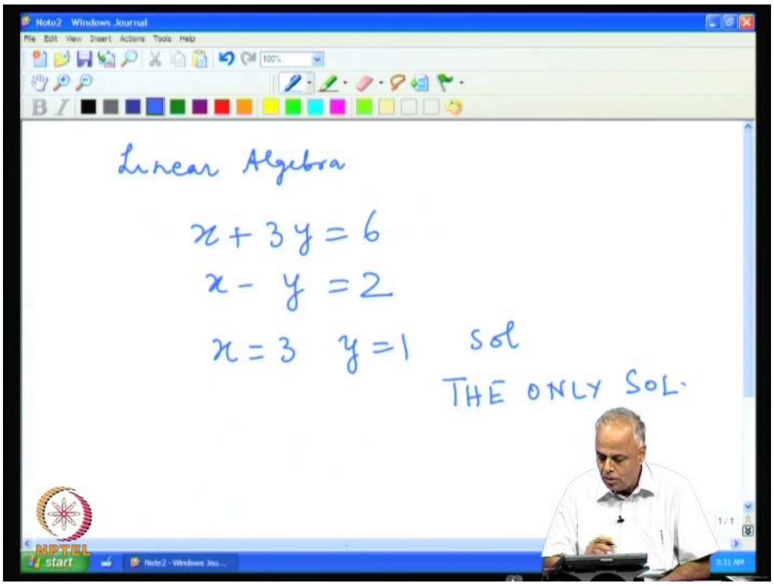
Prof. R. Vittal Rao
Centre for Electronics Design and Technology
Indian Institute of Science, Bangalore

Lecture No # 01

Prologue - Part 1

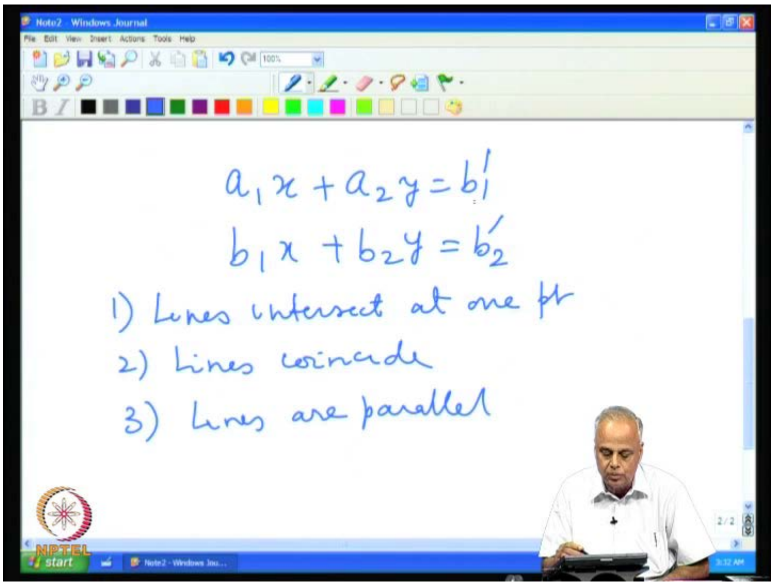
The subject of linear algebra has its origins in the study of systems of linear equations. Therefore, it is natural that we begin our course with the discussion on linear systems of equations. In the beginning, we shall keep our discussions not very rigorous, but only to motivate the basic and fundamental questions that we will be discussing in this course.

(Refer Slide Time: 00:56)



Let us begin with the simple system of linear equations; consider the system of linear equation x plus 3 y equal to 6 and x minus y equal to 2; it is very easy to see that x equal to 3 and y equal to 1 is a solution of the system; and more importantly, it is the only solution for the system. We can also geometrically interpret this as the intersection point of two lines; one of which is represented by the equation x plus 3 y equal to 6, the other one is represented by the equation x minus y equal to 2 and 3 1, are the coordinates of the points of intersection of these two points.

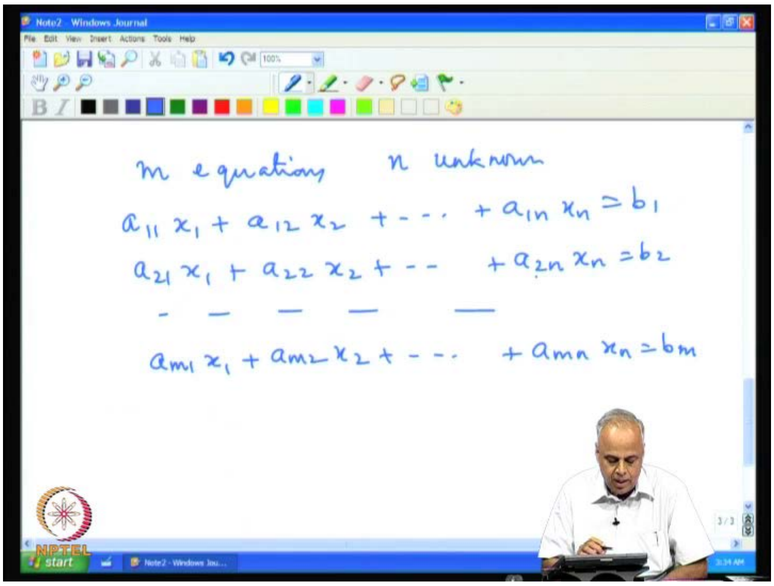
(Refer Slide Time: 01:51)



Looked at, from this point of you, we in general can consider a system of two equations, in two unknowns, and think of them as the intersection point of these two lines; however, we see therefore, that this could lead to many situations; one, the lines intersect at one point or the two lines coincide or the two lines are parallel. Therefore, when we have a pair of two lines, we can have these three situations. In the case one, when the lines intersected at one point, we see that there is a unique solution for the system given by the coordinates of the point of intersection.

When the lines coincide, we see that there are infinite numbers of solutions for the system, because every point represents a solution of the system. When the lines are parallel, we know that they do not need and therefore, there is no solution for the system. Thus, we will see that we could end up with several situations namely, lines intersecting at one point, lines inter coinciding with each other and therefore, giving infinite number of solutions or lines being parallel therefore, there is no solution; and therefore, in general, a system of two equations in two unknowns can have exactly one solution or can have an infinite number of solutions or can have no solution what so ever. Now it is clear, from this discussion that we could also have a situation with no solutions.

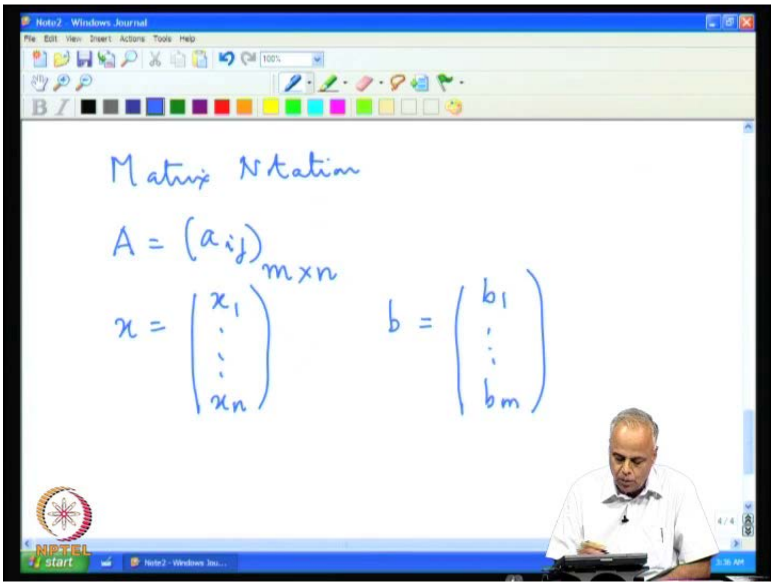
(Refer Slide Time: 03:54)



In general, we would be interested in considering m equations in n unknowns, these generally written if, the notation a 1 1 x 1 plus a 1 2 x 2 plus a 1 n x n equal to b 1, a 2 1 x 1 plus a 2 2 x 2 plus etcetera plus a 2 n x n equal to b 2 and so on. And the last equation is a m 1 x 1 plus a m 2 x two plus etcetera plus a m n x n equal to b m. Here, in the x the first index for example, in a 2 2 or in a m 2 or in a m n the first index refers to the number of the equation.

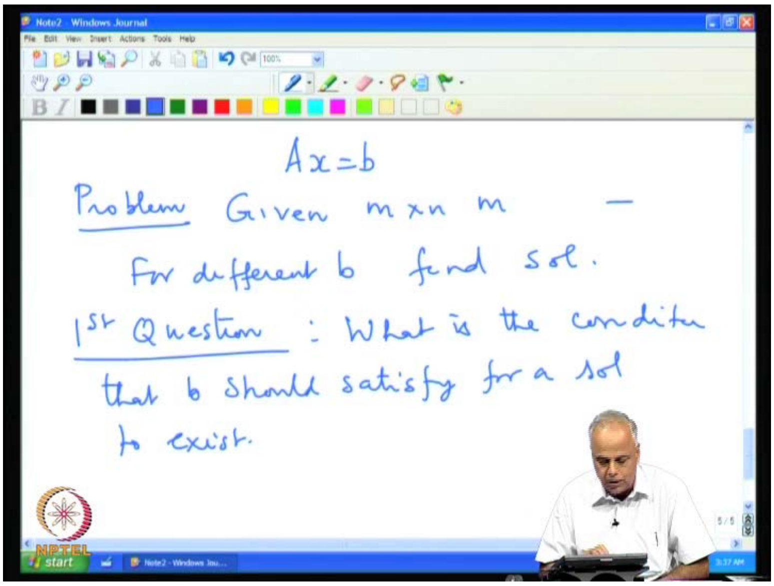
In the second equation, all of them will have the first index has 2 a 2 1 a 2 2 a 2 n and the second index for the x refers to which unknown they are referring to a 1 1 x 1 means, it is the coefficient of the first unknown a 2 2 x 2 means is the coefficient of the second unknown a m n x n means it is the coefficient of the n th unknown. Such, a system is what we will be considering in general.

(Refer Slide Time: 05:30)



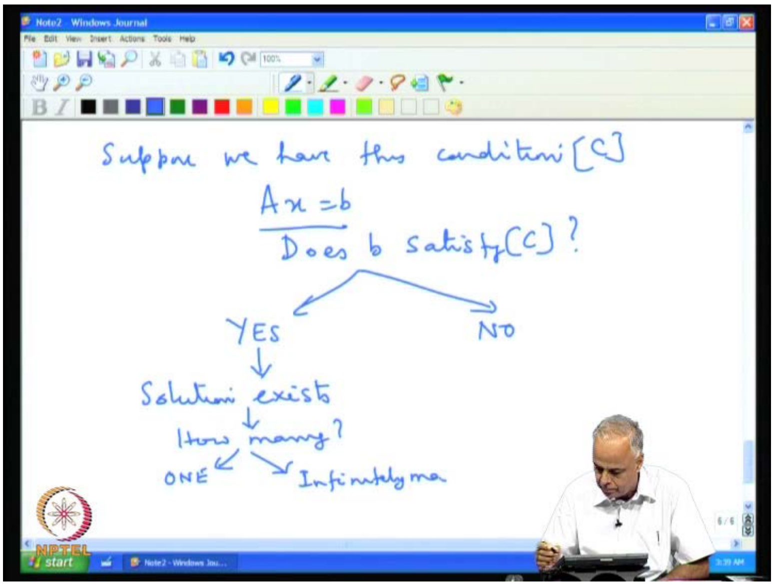
A system of this type, is written in a compact form in matrix notation as follows, we denote by A the matrix which is m rows and n columns with the entries as a i j. So, a i j is the coefficient coming from the i th equation corresponding to the j th unknown. We denote by x the single column vector matrix, which refers to the unknowns x 1 x 2 x n and by b the single column matrix b 1 b 2 b m, which corresponds to the right hand side of this equation. With this notation; in matrix notation, we can write the system as A x equal to b.

(Refer Slide Time: 06:24)



In general, therefore, we are interested in the system A x equal to b, what is the main problem? The main problem is given m by n matrix this is given. Then, for different b find solution, that is our fast fundamental problem of linear systems of equations. We have seen that, even when choose 2 by 2 systems, we could have a situation and there are no solutions and therefore, the first main question, that we have to answer is, what is the condition that b should satisfy for a solution to exist? This is one of the most fundamental existence questions, given the system A x equal to b; that means, the given the matrix A for what values of b will be have the solution for the system A x equal to b.

(Refer Slide Time: 08:16)



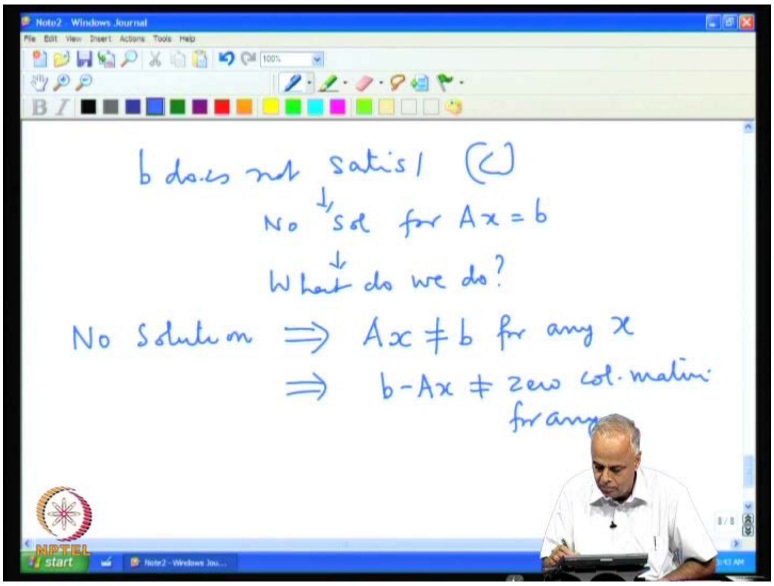
In other words, what condition should be satisfy in order the system a x equal to b has a solution? Now, therefore, given the system the first question that we will ask is, what is this criterion for existence? Suppose, we have found this suppose we have the condition c these conditions c then, given any problem any be any problem like this, A x equal to b the first, that we will ask does b satisfy c. Now, obviously, there are two possible answers for this, we may guess and answer yes and we may get an answer no.

We will let us discuss what we can do in both these situations. When the answer is yes, what is the conclusion that we can make? We can say, solution exist. When we say solution, you mean refer to the system A x equal to b. So, we know that the solution exist, then the fundamental question is, how many solutions exists? These questions arises because, we have seen in the two by two system there are situations, where we can get exactly one solution or situations where we can get infinitely many solutions. So, the answers are going to be one or infinitely many.

The question then is when do we get one solution under, what conditions do we get one solution? Then, we had these answers one solution infinitely many and the question when it is one? And what conditions we get one solutions? Under, what conditions we get infinitely many solutions? Now, having answer this question when there is only one solution, we want to find the solution; the unique solution which we know exists we have to find. Then, there are infinitely many solutions, what is the structure of all these solutions? That is the next question that you will have to answer. Now, having what this structure, we have a problem of plenty of the situations. There are many solutions, is their some way of choosing one solution.

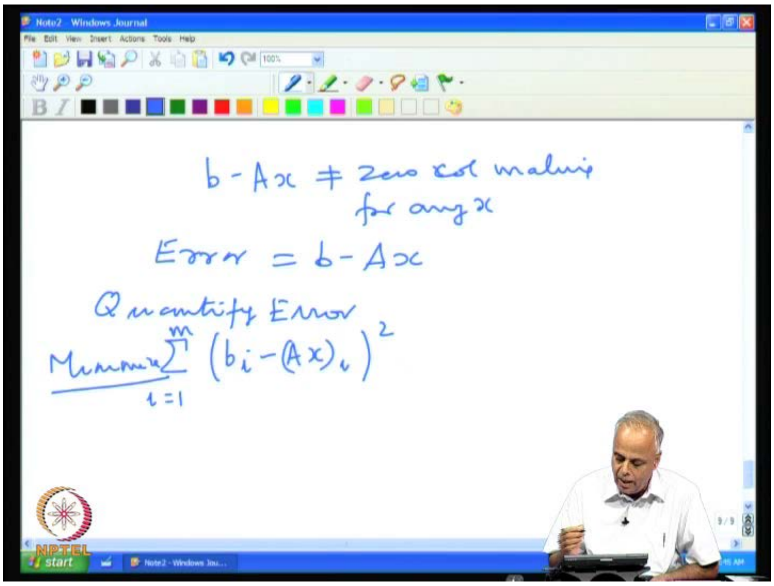
Criterion to choose, a unique represented among all these solutions we may choose a unique representative. Then finally, find the unique represented therefore, when we are a situation where these satisfies the consistency condition. Solutions are guaranteed then, we are in several situations, one solution are infinitely many solution. In the case, have one solution, you would like eventually find the solution; in the case of infinitely many solutions, we would like the find some criterion by which we choose a unique representative solution and get that representative solution.

(Refer Slide Time: 11:53)



Now the problem is, what happens b does not satisfy the criterion c, in that case when b does not satisfy the criterion c then, what can we conclude? We can only conclude no solution for A x equal to b, because c was the criterion for a solution to exist and b does not satisfy it and therefore, we would like to conclude that, there cannot be any solution. If there is no solution, what do we do? For this, we would like to analyse, what exactly does; the fact that there is no solution imply. Therefore, what do we mean by no solution this means that, whatever x I choose and calculate A x is not going to be equal to b for any x that means, b minus A x is not going to be the zero column matrix for any x.

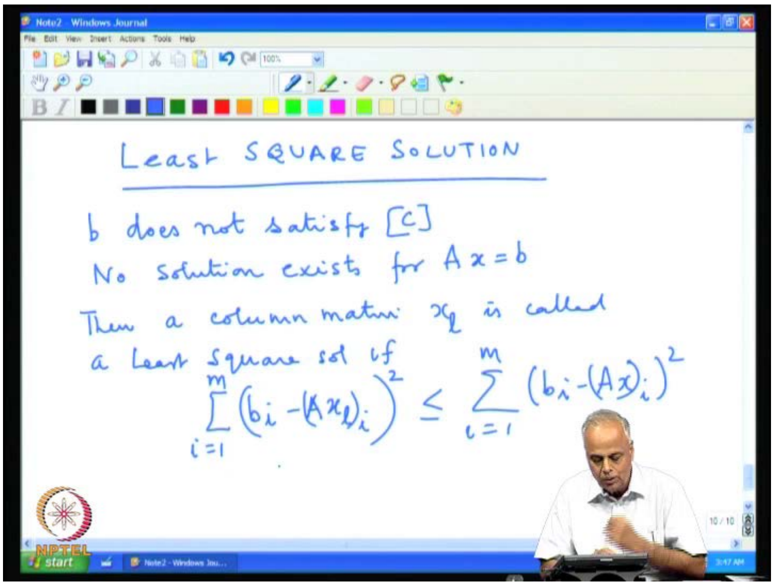
(Refer Slide Time: 13:33)



This means, if I look at b minus A x, I would have like, this to be zero if x where a solution, but since x is cannot be solution for any x b minus x is not equal to zero vector zero column matrix for any x and therefore, if we take x as a solution there is going to be an error. That error is given by this difference b minus A x. Now, what is the general idea when you get an error? The general idea when you get error is, would like to minimise that error. How do I minimise this error? We quantify this error first.

The error is in the form of a column matrix, we now quantify this error by n number, how do we do this? We look at the i th coordinate of the error that is a number, take this square of that and now do this for every coordinate from i equal to 1 to m and add this; this gives us the square error, from each term added together in the entry of the error vector. Now, we would like to minimise this quantified error. So, can we do it? We shall see later that this is possible and this leads as to the notion of the least square solution.

(Refer Slide Time: 15:26)



What do we mean by a least square solution? So, suppose b does not satisfy the consistency condition c. No solution exists for A x equal to b then, a column matrix x we will call it l to denote the fact, that it is going to be a least square solution is called a least square solution. If, minimises the squares of the error what does it mean? If we now look at the error obtained by taking this x and calculate the error. In each entry in the error matrix square it and add all the errors, that error should be the least among all possible errors, that is if I take any vector a x and calculate the error it cannot be less than this. So, the least squared solution is a solution, which minimises this particular quantification of the error. Obviously, therefore, different notions of errors quantification could lead as to different types of error solutions, but we will be concentrating only on these least square solution.

Therefore, in the case of, we can always we can show that, we always can get a least square solutions if b does not satisfy c. We can show that, there is a least square solution and therefore, let us get back to the situation the next question therefore, we would ask is how many least square solution? Now get into a loop again, just as in the case when b satisfied c there was a solution and now, we does not satisfies c there is a least square solution. In the case b satisfied how many? We ask, there are one or infinitely many. In these case also, we get either only square solution or infinitely many least squared solution.

Now therefore, the question is again, when does this happen? And when does this happen? How many answer that the question, that you would like to answer is finally, find the least square solution. In the case of the infinitely many remember, that in the case of infinitely many solutions when b satisfied c we said, we want the criterion for choosing a representative. Again here, you would ask for choosing a representative criterion for a representative because, we are a problem of plain t we have too many least square solutions eventually we have to pick one solution. So, we need some criterion for picture having got that criterion, I would final goal should be find the representative least square solution.

Therefore, while dealing with a linear system of equations, there are several questions that arise. Now, over one of our major goals will be to get the answers to all these questions and look at the generalisations of all these questions that are possible and we will develop the correct mathematical framework in which, we should ask this questions. Now let us therefore, understand that a linear system of equation is going to involve several questions which require fundamental answers and fundamental questions which will require the theory to get the answers, which would also required techniques to get the solution.

Let us, next get back to the question of, how do we handle all these question? What is the natural way to go about finding the answer? The natural way that the mathematician does is whenever have a problem, which is typical to solve look at, the simplest version of the problem and see what are the mechanisms involved in the simplest version and that knowledge gain from that is going to help him to handle the most difficult or the most general situation. So, the first thing that you would like to ask is what are easy systems? The systems which can be easily solved which for, which all the answers can be given without much hard work.

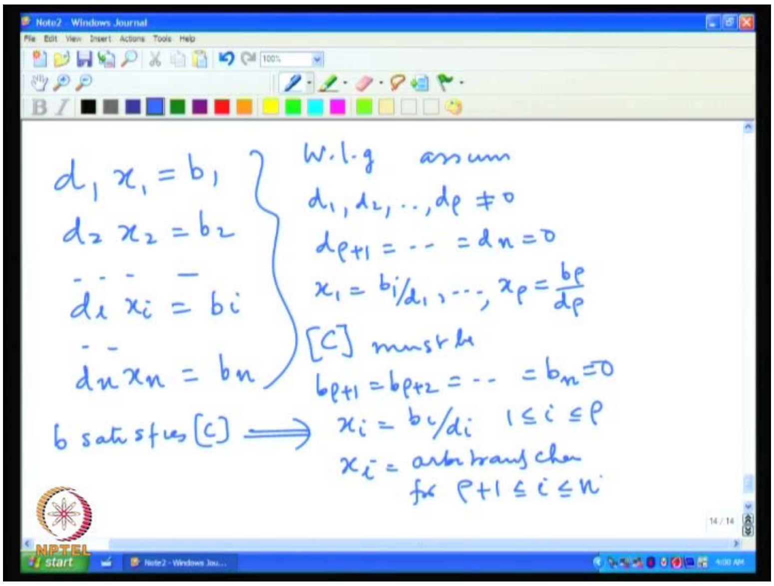
To answer this question, you must first understand why are the general systems difficult? Why are the general systems difficult? The difficulty lies in the fact, that each equation involves every one of the variables or probably each equation will involve several variables and therefore, no equation on it is own is able to help you, to determine one of the unknowns. This is what we know, call us the variables or the unknowns are all coupled in a general system and this is what makes, what do we mean by solving the system? We want to get x one separately; we want to get x two separately; we want to get x three separately; x four separately and so on.

That means, we want to uncouple them that is essentially, what is meant by solving a system? You want to uncouple all the variables and therefore, it is this uncoupling process that looks very complicated because, the coupling more the coupling; the more number of variables. Each one of these equations involves, the more complicated the uncoupling is going to b therefore; let us first look at a system, where such an uncoupling is easy.

A system where uncoupling is easy, what is this system? Obviously, do you know coupling at All right, in the beginning? In this equation cell there is no coupling then, that is any easy system because, the work is already done. That is a system where there is no coupling, what is such a system? Such a system each equation involves only one of the unknowns. So, each equation involves only one unknown. Let us look at the simple situation, where the first equation involves the first unknown; second equation involves the second unknown; the nth n th equation involves m n.

Let us even make it clear that, there are n unknowns we are n equations and i th equation involves such i th unknowns. So, n equations n unknowns i th equation involves i th unknown x i. How does such a system look like? It involve look for, the i th equation will involve only the i th unknown.

(Refer Slide Time: 26:02)



We will have the first equation will be d 1 x 1 equal to b 1 d 1 x 2 equal to b2. Note the first equation was only x 1; the second involves only x 2; the ith equation involves only x i and so on and the nth equation involves x. So, this is one of the simplest situations of a system. We have n equations; we have n unknowns; the variables are the unknowns evolve completely decoupled and i th equation involves only i th unknown. Therefore, we can easily solve, try to solve at least these equations.

Let us look at, how we would go about solving it? If you want to solve the i th equation, we would like to write x i equal to b i by d i, but then we cannot divide by d i or if d i is 0. So, the equations which are where the d i are 0 are going to be our problematic. So, let us separate this how? Let us without loss of generality assume, the first row of the d are not 0 and the remaining d are all 0.

In that case, at least we get from the first equation x 1 equal to b 1 by b 1 and so on and the rho th equation gives x rho equal to b rho by d, but now we will look at the rho plus one th equation, the left hand side for the equation will be 0. Whereas, the right hand side will be b rho plus one and therefore, a solution will exist then, both the sides will be zero only and therefore, b rho plus 1 must be equal to 0 continuing this process, we get the consistency condition must be b rho plus 1 b rho plus 2 etcetera b n must all be 0.

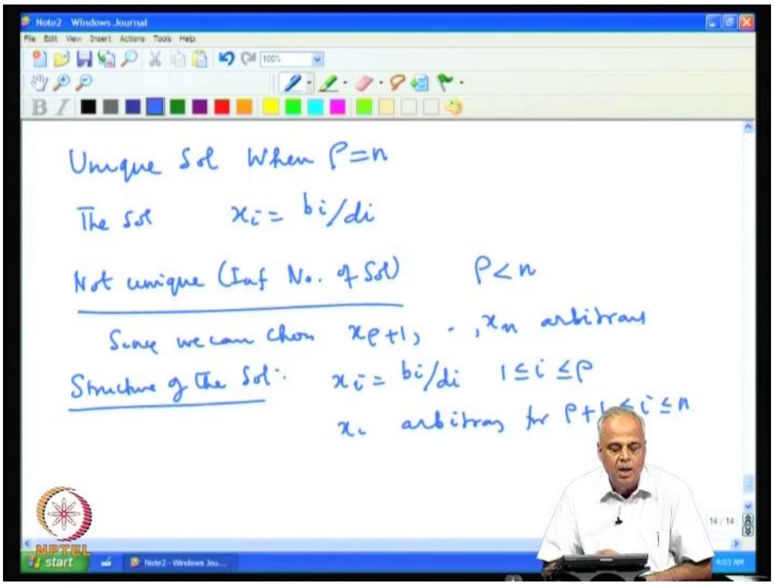
Therefore, a system like this will have a solution if and only if the rho whenever, thus the 0 on the left hand side thus, the corresponding these are also 0. So, we are assuming that the left hand sides are all 0 from the rho plus 1 th equation onwards and therefore, the b must all be 0 from the rho plus 1 th entry onwards. So, therefore, the consistency condition for such a system becomes b rho plus 1 b rho plus 2 b n are all 0.

Now, what do we get from this? If b satisfies this condition, what does that imply? Then, the first rho were unknowns x 1 is b 1 by d 1 x 2 is b 2 by d 2 x 0 is b rho by d rho. These are got and the remaining variables whatever, value I choose the equation only give zero equal to zero zero equal to zero zero equal to zero. So, what we get is x i are given as b i by d i for one less than or equal to i less than equal to rho and x i is arbitrarily chosen can be arbitrarily chosen for rho plus 1 less than or equal to i less than or equal to n.

Now, remember the first question, we ask was what is the consistency condition or what are the conditions? Under with the system has a solution. For that, now we have an answer the conditions are whenever 0 d rho plus 1 d n are 0, the corresponding d rho plus 1 b n must be 0, that is the consistency condition. Now, when that condition is known given a b we ask thus b satisfy this condition. Suppose, we satisfies this condition then, we have the answer that the solution must be have the form x i equal to b i by d i for one less than or equal to i less than or equal to rho and x i equal to any arbitrary value for i between rho plus one and n.

Now, this is there are lots of arbitrary value that we can choose for x i from rho plus 1 onwards therefore, even if rho is n minus 1 then, there is an i for which x i can takes several values and therefore, we will get an infinite number of solutions and when will therefore, we get a unique solution. When all the exercises are uniquely determine, that will happen then rho is equal to n. So, what is our conclusion out of these, we get the unique solution when rho equal to n.

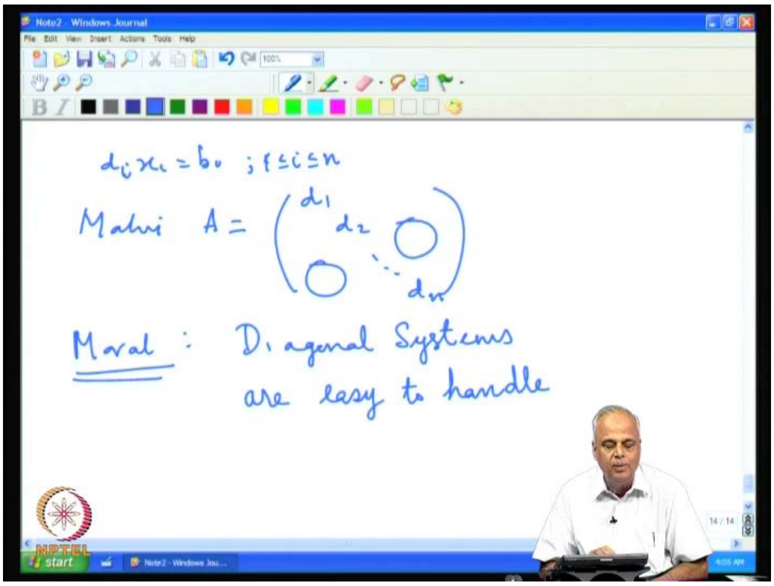
(Refer Slide Time: 32:00)



Remember, the question we ask, when does it happen unique solution? We have the answer for this, when rho equal to n and then, what is the solution? And the solution is x i equal to b i by d i. Thus, in this case, in the simple system at least then rho equal to n we have all the answers. When is the solution not unique? That is infinite number of solution then, rho is less than n. Since, we can choose x rho plus one x n arbitrary.

What is the structure of the solution? The structure of the solution is x i equal to b i by d i for 1 less than or equal to i less than or equal to rho x i arbitrary for rho plus 1 less than or equal to i less than or equal to n. Now, we have not answer one further question. Whenever, we have infinite number of solution; we would like to have a criterion by which we choose one solution as the representative solution. We will come back to this question little later is not very difficult, but we will handle them all together.

(Refer Slide Time: 33:50)

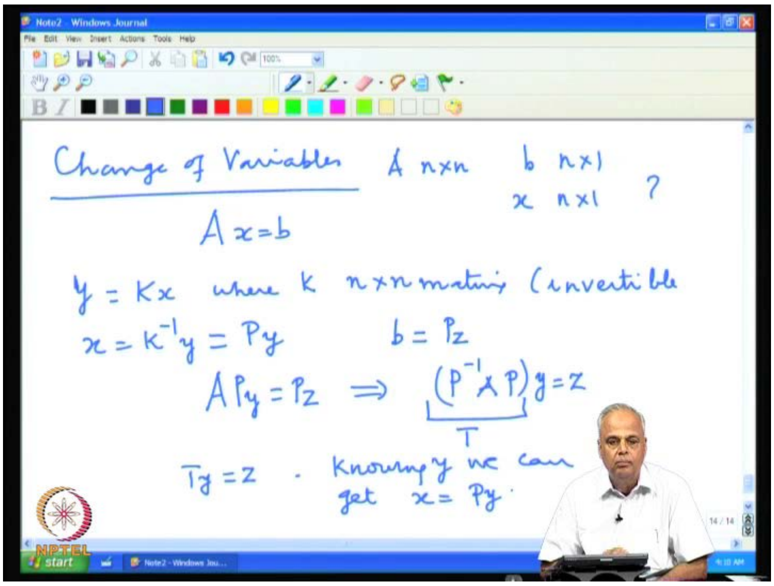


But, the model of the story of discussion of this simple problem is that, when we have a system d i x i equal to b i 1 less than or equal to i less than or equal to n. A system of n equations in n unknowns, where the variables have all uncoupled most of our answers can be easily caught. Now, what is the structure of the system? The important thing in the system is the matrix A for such a system is a diagonal matrix because, a 1 1 is the coefficient of the x 1 and that is d 1 and remaining coefficients in the first equation are all 0.

Similarly, the second coefficient in the second equation is 2 and all others d 2 and all others are 0 and so on and therefore, the coefficient matrix is a diagonal matrix. The moral of our discussion for such special easy system is diagonal systems are easy to handle. Let us put it that therefore, the diagonal system are easy to handle. We will discussing this one more time in greater detail in more rigorous situations and more rigorous language, but at the simple motivation. We have seen that when system is diagonal, it is easy to handle almost, all the questions come out; the answers come out fairly in straight forward.

Now, therefore, what should be our next move? The idea always is, if we know how to solve or how to handle one particular problem? In next level complex problem, would be attempted to be brought down to this easy level; given, a general system n by n system. Let us, take to n equations in unknowns the easier version. So, given a general n by n system A x equal to b can be reduce it to the analysis of a diagonal system. can be reduces to the analyses of a diagonal system

(Refer Slide Time: 37:11)



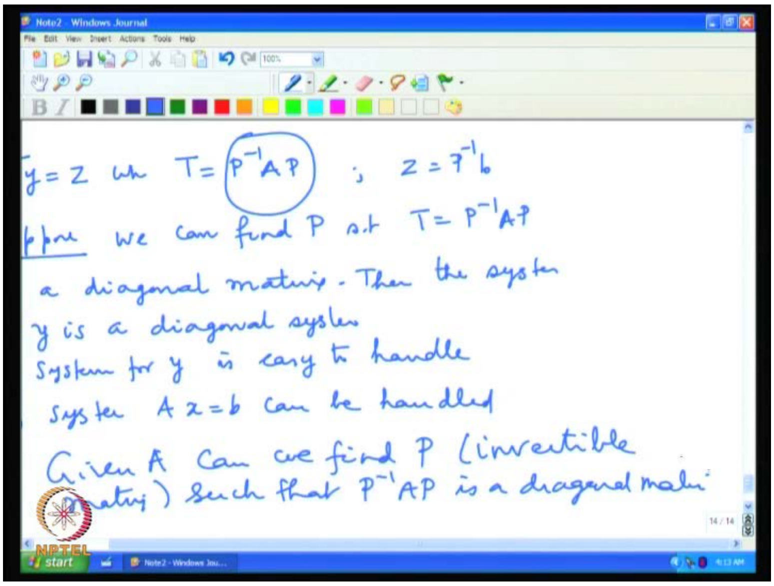
What do we mean by this or how can we go about getting to this diagonal system? This where, we like to introduce then the idea of change of variables. Suppose, we have the system A x equal to b, A is an n by n matrix, b is an n by 1 column vector, x is an n by 1 column vector. We will just be found, we want to know, what is x for that A x equal to b. Now, we introduce a new variable y which is equal to say some k times x, where k is an n by n matrix now what the idea? The idea is somewhat try to get an equation to y, which is easy to solve.

What is the use? The idea is in y know why? I can find an x, how do I know x from y? That I would be able to write x is equal to k inverse y provided k is an invertible matrix. We will say, k is an n by n invertible matrix. So, we introduce a change of variables y equal to k x where, k is a invertible n by n matrix then, x becomes k inverse y or in a standard most of this, books will use the notation p y.

So, we will say x is equal to p y. Now, therefore, b also we can changes as some p z. We have introduced to change of variables y will be the new unknown and z will be the new right hand side new known. What happens to the system? The system becomes A times x, x is P y equal to P z because b is equal to P z. Now, P is invertible because k P is k inverse and k is invertible. This implies P inverse A P y is equal to z therefore; the equation for y becomes P inverse A P y equal to z.

We will call P inverse A P y assume T. So, that becomes T y equal to z. If we know y then, we get x as x is equal to k inverse y. So, therefore, knowing y we can get x as P y. The problem therefore, is can we solve y. We can solve y if T were a simple matrix. What do we mean by a simple matrix? Namely, the diagonal matrix.

(Refer Slide Time: 40:55)



Suppose, we can make d diagonal, we have T y equal to z where, T equal to P inverse a P and z is equal to P inverse b or b is equal to b z. Suppose, we can make find P such that, T equal to P inverse A P is a diagonal matrix. Then, the system for y is a diagonal system and we have seen the diagonal systems are easy to handle. So, system for y is easy to handle and since, we the knowledge of y; we can get x. It says, the system the given system A x equal to b can be handled.

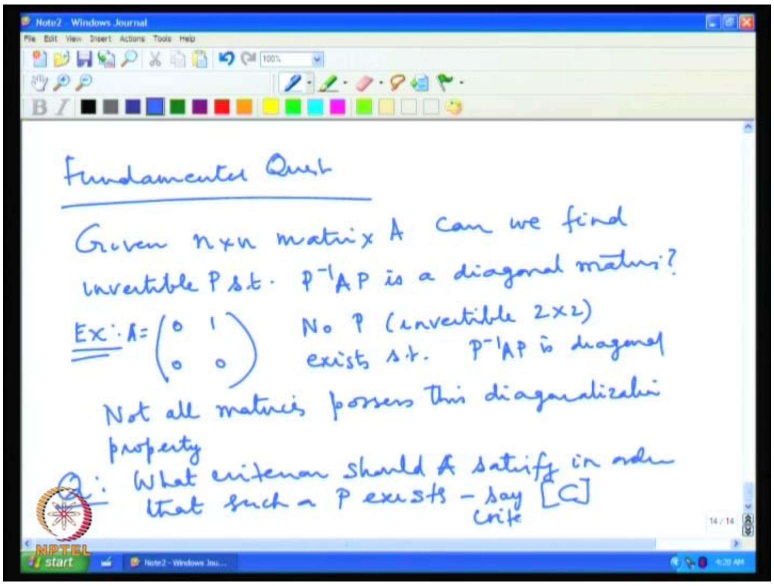
Therefore, the fundamental question that we has made is if, this passage of change of variables where then, any system of n equations in n unknowns can be converted to a diagonal system. Therefore, the fundamental question is does this sort of change of variable work. So, given a matrix A can we make, the matrix into a diagonal matrix by a transformation of the form P inverse A P.

Given A, can we find P an invertible matrix such that, P inverse A P is a diagonal matrix. This is another fundamental question in linear algebra this called a diagonalization process of a matrix and we will look at the answer to this question, as we go along, as one of the most fundamental problems in linear algebra our diagonalization of this matrices.

Now, we shall see with few examples, but these diagonalization question creates certain problems namely, that not all matrices can be diagonalized and hence, we get into a fundamental question that, this process of converting a general system of equations to have diagonal system may not always work. It tends out for certain class of matrices; certain types of matrices, we can find of P. So, that P inverse A P is diagonal and for certain types of matrices we cannot find A P so, that p inverse a p is diagonal .

Now, therefore, it raises the question under, what conditions can a matrix A provide you with the P such that, P inverse A P is a diagonal? We have know, one more criterion to be form. Remember, just as we said, what is the condition we should satisfy in order that A x equal to b as a solution? Analogously now, we are asking what is a criterion that A should satisfy in order that P inverse A P is a diagonal matrix. Consequently, the moment we find this criterion; we will again ask a series of questions namely, given a matrix A check whether, it is satisfy this criterion to be does.

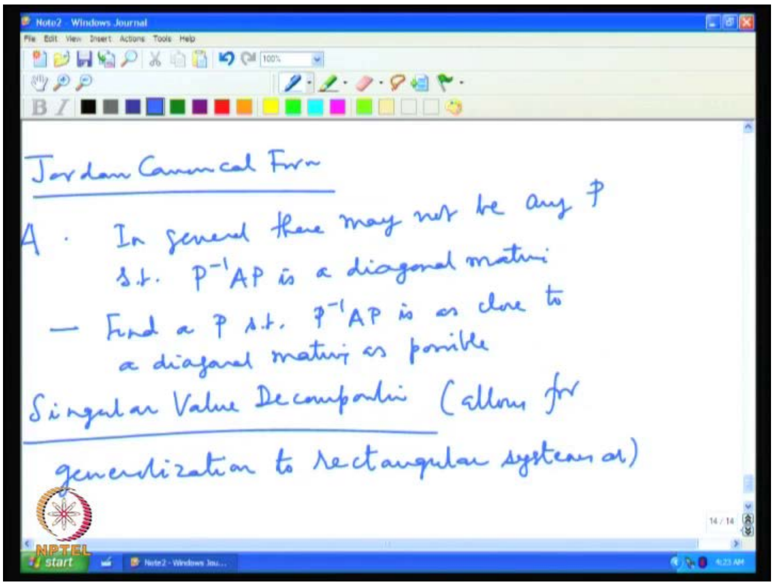
(Refer Slide Time: 45:53)



The fundamental question therefore, is given n by n matrix A can we find invertible P such that, P inverse A P is a diagonal matrix. The answer to this question is very complex and this leads to a lot of very interesting canonical forms of matrices and we shall be discussing all of them at least, several of them. The first think is to look at one example, if you look at this matrix, we leave it as an exercise to verify that no P which is invertible 2 by 2 exists such that, P inverse A P is diagonal and therefore, we are into certain difficulties that not all matrices possess this diagonalization property.

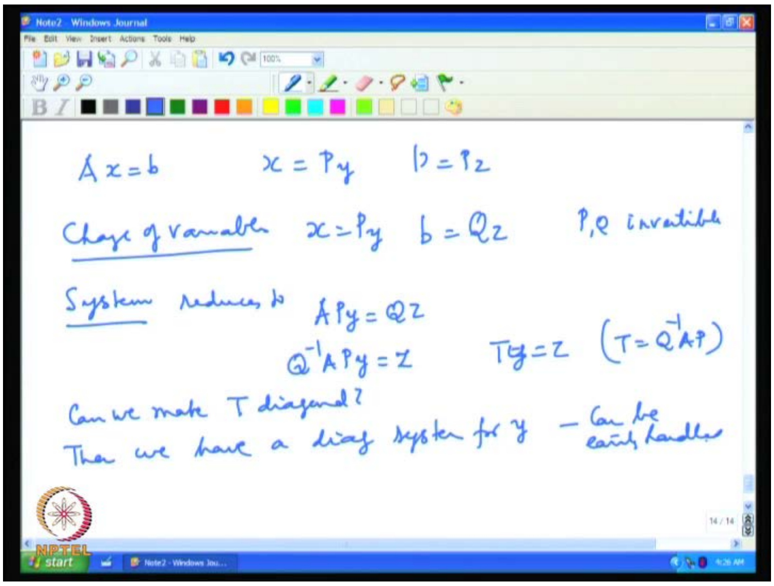
Therefore, the fundamental question is what criterion should A satisfy in order that such a P exists? Let us, call again such a criterion say criterion c. Now, having what this criterion the question therefore, would be, if the criterion is satisfied will be able to do this diagonalization; if we are the criterion is not satisfied there will not be able to do this diagonalization. Now, the diagonalization non possibility, not possible leads to two important things, which will be discussing in the class one is the so called Jordan canonical form and the other, which is very useful in applications called the singular value decomposition.

(Refer Slide Time: 49:28)



Both these are very important, what is this Jordan canonical form? It is we have A, in general there may not be any P such that P inverse A P is a diagonal matrix. So, what do we do? So, I would say if I have to find a P, such that P inverse A P is as close to a diagonal matrix as possible. Find a P such that, P inverse A P is as close to a diagonal matrix as possible. Now, what do we mean by as close? Now, making these things formal, precise leads you to a notion of Jordan canonical form. What is this singular value decomposition? This is very important for us, because so far, we have been discussing a system of equation, where we have n unknowns and n equations, but when we look at the singular value decomposition. We will see that this also now allows us to generalise this notion of singular value decomposition to rectangular systems, where the number of equations may not be equal to the number of unknowns. So, this is allows for generalization to rectangular systems.

(Refer Slide Time: 51:59)



What is this singular value decomposition? Now recall that, we had the system A x equal to b, and we introduce the transformation of variable x equal to P y b equal to P z. Here, we are use the same matrix P to transform x as well as y. Suppose, we use when change of variables x equal to P y and b equal to Q z where P Q are invertible matrices. Then, the system reduces to A P y equal to Q z or Q inverse A P y equal to z. So, if we now call Q inverse A P as T this becomes T y equal to z T equals Q inverse A P. We can ask, can we make T diagonal? That is, now we are asking not the same P to be use for both x n be transformations, but we use a different transformation for b and different transformation for x and then try to convert the equation in terms of this new variable and you by chance, we can find a Q and P. So that Q inverse A P is diagonal then, we have a then diagonal system for y and therefore, can be easily handled.

And one, we are a diagonal system for y and the handled, we are our matrix got x because, x is equal to P y now the answer to this question that can be it make the diagonal. The answer is yes, given any matrix, because reason of the conditions on A we will put the right framework, but generally we can say, the answer is yes and therefore, this gives as a better handle on the system by using two different coordinates; one to transfer the unknown x, the other one to transfer the known b. Now you see that the transformation for b involve the matrix Q. Now, in the case of a rectangular system, b will be m by 1 so, we can choose Q to be an m by m matrix and x was a will be n by 1. We can choose p to be an n by n matrix.

So, behave one transformation again for the unknown and one transformation for the known b and then, we would ask a question whether Q inverse A P is almost. We cannot ask, whether it is to be a diagonal because now, since A is rectangular Q is m by m p is n by n. The resultant will be an m by n matrix, we cannot ask talk about, a diagonal rectangular matrix; we would like to ask something like almost diagonal. The idea of what is almost diagonal means leads to the notion of singular value decomposition. So, therefore, summing up what we have is? That we have a system of equation, which we can analyse by reducing it to the simplest system and simplest systems are the diagonal systems in the case of, the square systems, where we have number of unknowns equal to the number of equations, then we can do two things.

One is, we can reduced system to a diagonal system by using two different transformations; one for the unknown x and one for the known b; but if you want the same transformation for both b and x, then we may not be able to diagonalize; we may nearly diagonalize it, and then try to analyse system. This leads to the two notions, these svd or the singular value decomposition and the other one the Jordan canonical form. Now, this gives us roughly, the idea of the problems that we will be handling in this course. The main theme will be solving systems of equations; this would amount to studying the structure of a matrix, which would then lead us to generalisations of further linear transformations on vector spaces and all; but the motivation comes from solving a system of the equation and how to reduce it to a simple. .