Advanced Matrix Theory and Linear Algebra for Engineers

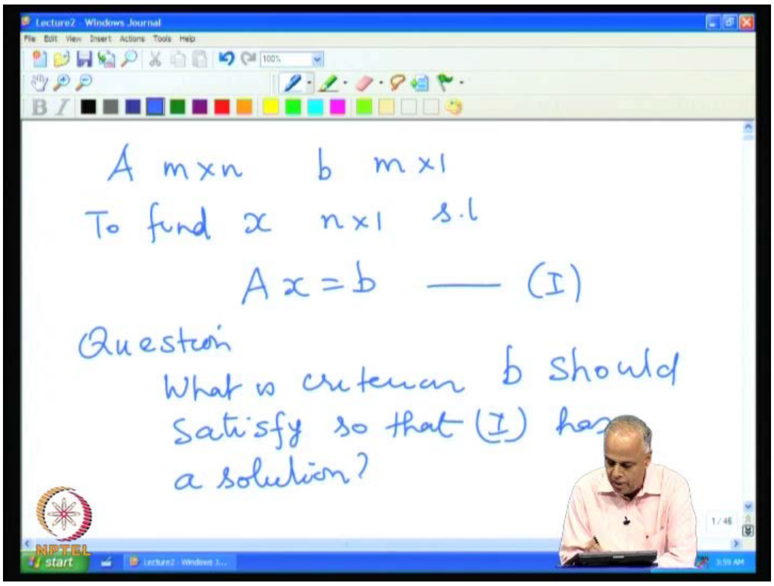
Prof. R.Vittal Rao
Center for Electronics Design and Technology
Indian Institute of Science, Bangalore

Lecture No. # 02

Prologue-Part 2

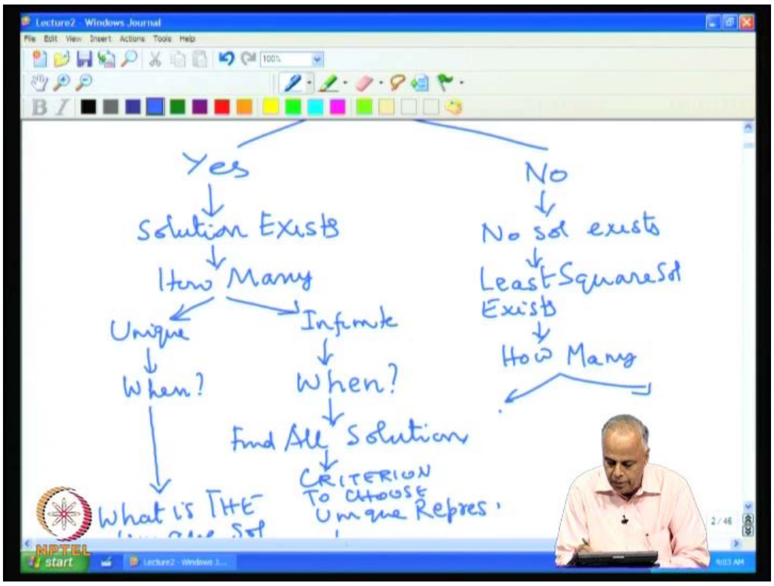
In the last lecture, we discussed two important problems in Linear Algebra, and we shall review this since, these forms a basic foundation for many of the things that we are going to study.

(Refer Slide Time: 00:32)



The first problem we studied was linear systems of equations. We have an m by n matrix and we have given an m by 1 matrix, and we want to find an x an n by 1 matrix, such that A x equal to b. This is the first fundamental problem that we looked at. The first basic question that we raised was, what is the criterion b should satisfy, so that the system one has a solution? That was the first fundamental question.

(Refer Slide Time: 01:38)

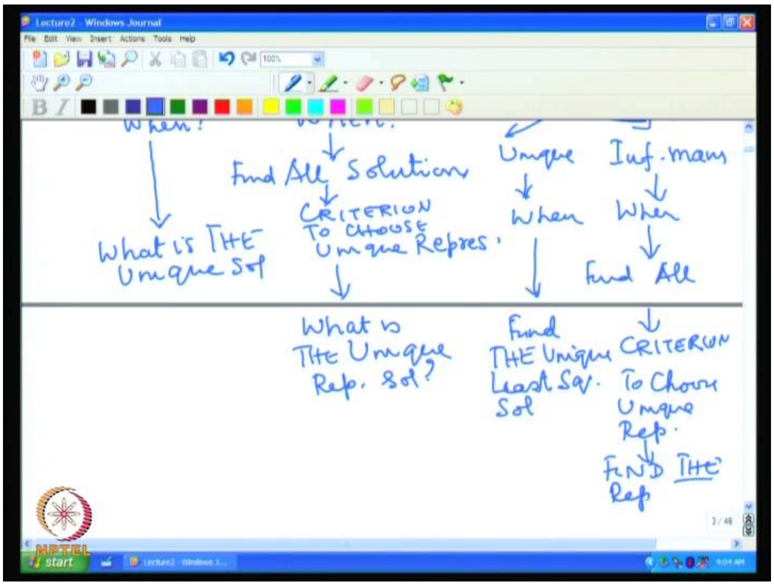


Once, we have the answer to this question. Given b, we ask, does b satisfy c? Naturally for this, we may get two different answers possibilities; we may encounter a situation; we get yes as the answer or we may encounter a situation where, we get no as the answer. Let us look at, what both cases lead to? When it is yes, the thing that we can conclude is that solution exists; and we say solution, we mean a solution to the system A x equal to b.

So, the moment we are sure that there is a solution; this leads to the following question. How many solutions are there? As we saw last time, this is two possible answers, a unique solution or infinitely many solutions. Now, naturally we ask when and what conditions we get a unique solution? And when and what conditions, do we get infinite number of solutions? And when we have the answer to this question, we naturally ask, what is the unique solution? When we are encountering a situation, where we have an infinite number of solutions, we would like to know all solutions. We must find all solutions; however, there is a problem of plenty. So, we must have some criterion to choose unique representative; among the solution, we must be able to choose a unique representative solution.

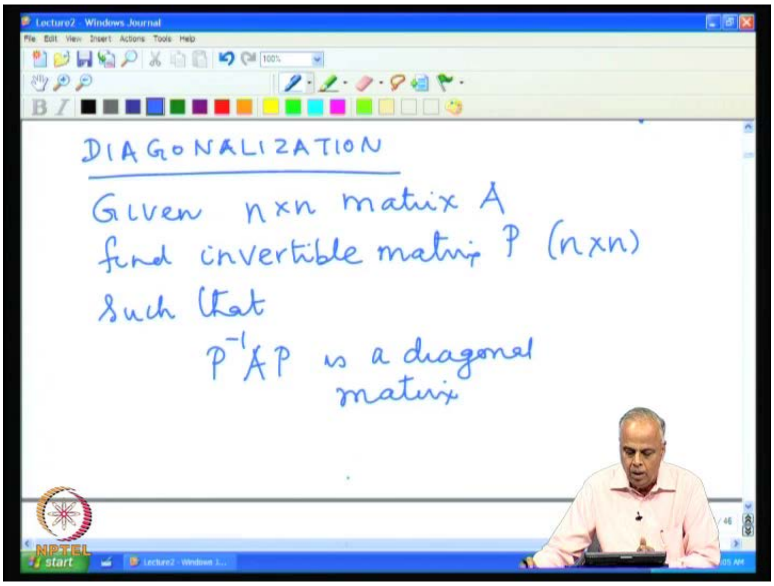
Now, once you have this criterion, the question we ask is what is the unique representative solution? Now let us, go back to the other situation, where we may encounter no as the solution? The answer to the question that is we does not satisfy the criterion for the existence of a solution. Then the conclusion we can make is no solution exists. Then we found last time that under these circumstances, we can find, what is meant by least square solutions? That is the solution that minimizes the error between b and A x. Then once we have assured that the d square solution exists, we again ask, how many least square solutions are there? How many least square solutions are there?

(Refer Slide Time: 04:51)



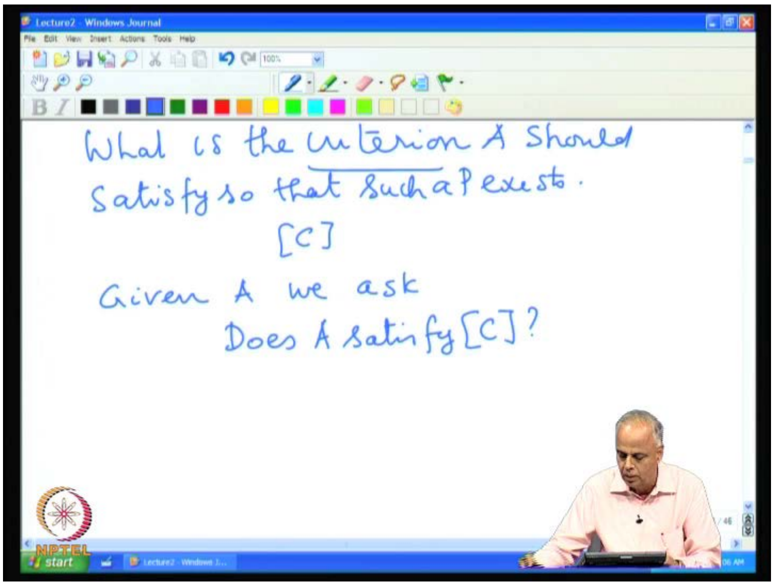
Again, we encounter the same situation as before, unique and infinitely many. Again, when is it unique? When is it infinitely many? Then, one shows the answers to these questions. We want to find, the unique least square solution and then, it is infinitely many and we know when it is, we want find all find all the least square solutions. Again as before, we want a criterion to choose criterion to choose unique representative. Then finally, find the representative. So, these are the fundamental questions that arise, when we are dealing with system of equations? Our aim will be to find the answers to all these questions.

(Refer Slide Time: 06:21)



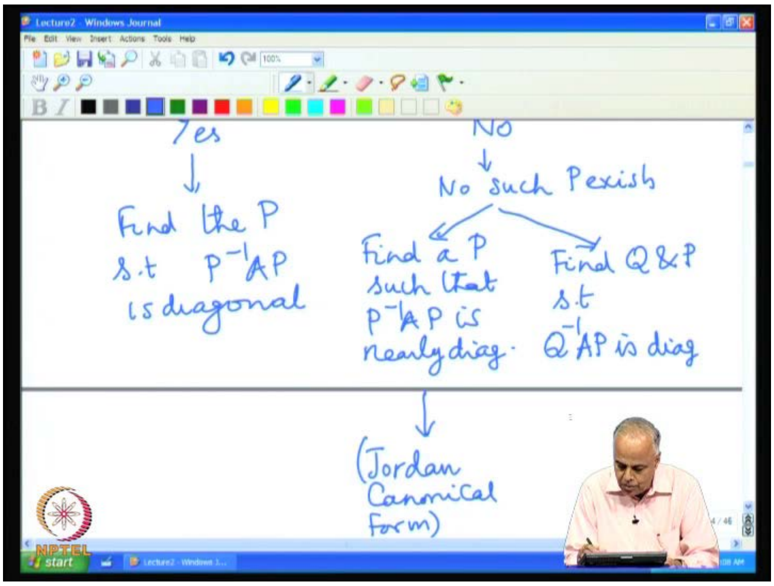
The other problem that we discussed was basically, about diagonalization. The question was given, an m by n matrix p given an m by n matrix A, find invertible. Now, let us look at a simple situation first; let us take a square matrix n by n square matrix and then, find an invertible matrix P, which is also n by n such that, P inverse A P is a diagonal matrix.

(Refer Slide Time: 07:30)



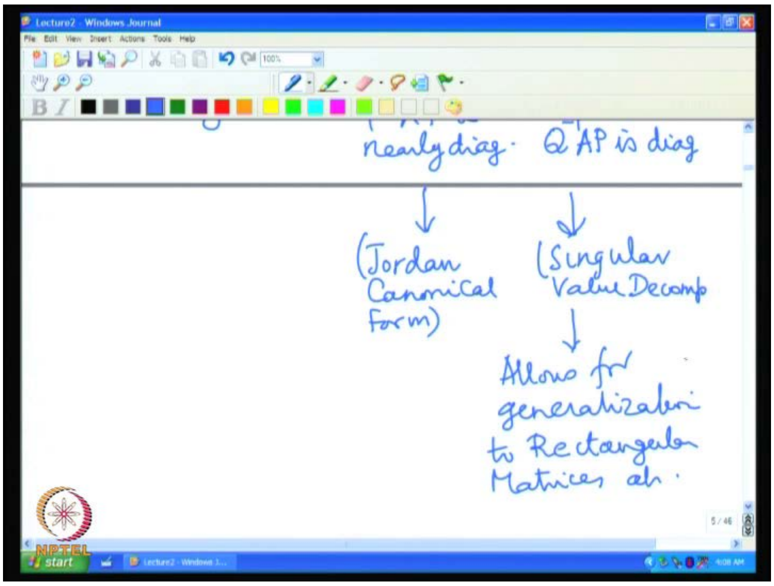
We found, that this is not always pass it well and therefore, we ask a question, what is the criterion A should satisfy? So, that such a P exists. Let us, called this criterion as C so, we will use the symbol C for the criterion C.

(Refer Slide Time: 08:04)



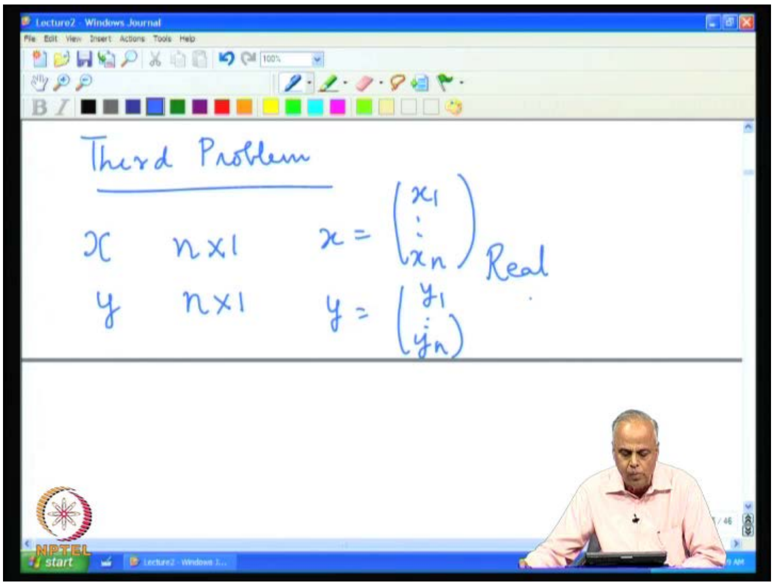
Then, the given A we ask, does A satisfy C? As before, they are lead to two possible answers yes and no. Again, when it is yes? We ask find the P such that P inverse A P is diagonal, when it is no? We know, there is no such P, so no such P exists. Now in this situation, there are two types of analysis we can do; the one is find a P such that, P inverse A P is nearly diagonal, can we do such a thing? And an answer to this, would lead as to what is known as the Jordon canonical form? On the other hand, there is another option available for as, is instead of finding one such P, find Q and P such that, Q inverse A P is diagonal.

(Refer Slide Time: 09:55)



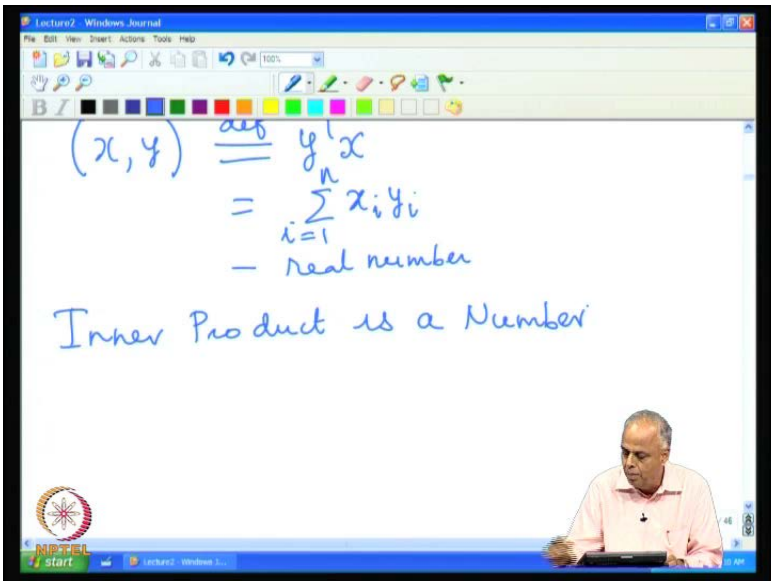
We shall see that, this leads to an answer to this leads to, what is known as the singular value decomposition of the matrix A? Now, the advantage of this singular value decomposition is, it allows for generalization to rectangular matrices as well. It allows for generalization to rectangular matrices. So, these are two basic problems which we discussed in the last lecture.

(Refer Slide Time: 10:47)



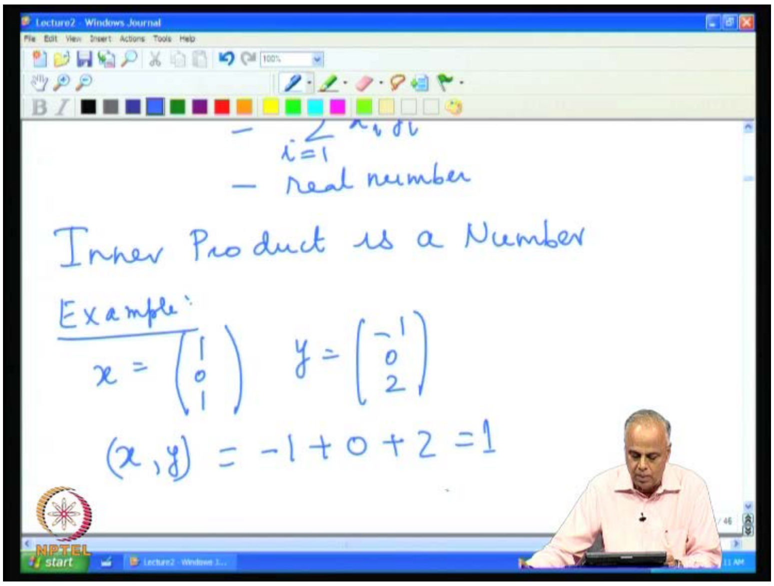
Now, we shall look at a third important problem. It must mention that, the other two problems are interrelated. You may recall the second problem a rows out of reducing the system to a simpler system. Let us now, look at another important problem which is as follows. Suppose, we start with a vector x are what we call, a column matrix and another column matrix say x is x 1 x 2 x n and y is y 1 y 2 y n and let us say all are real. So, let us say, we have two vectors or two column matrices each having n entries.

(Refer Slide Time: 11:42)



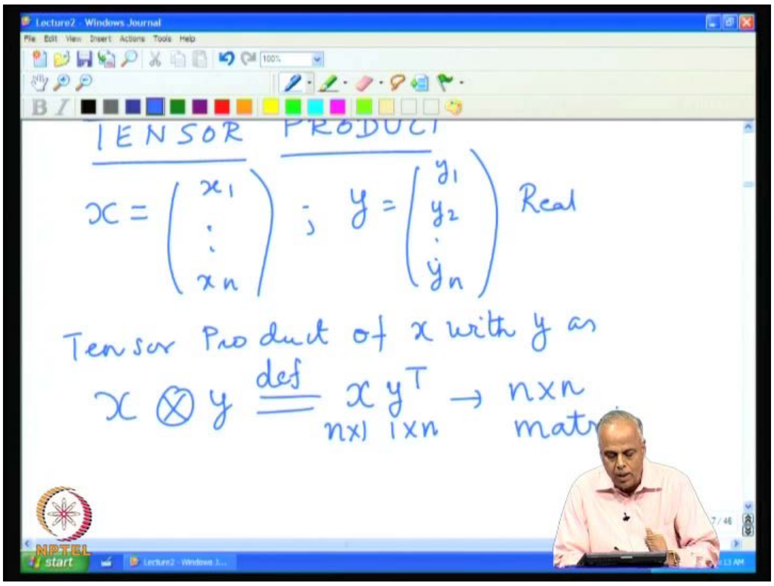
Then we define, what is known as the inner product of x and y? We denote it by x comma y both are bracket and we define, it as y transpose x which is simply summation i equal to 1 to n x i y i. This is a real number thus, the inner product of these two is a number; the inner product of two such column matrices is always a number.

(Refer Slide Time: 12:43)



Let us look at a very simple example; let us take x equal to 1 0 1 and y equal to minus 1 0 2 then, the inner product is simply the sum of the product of the corresponding components which is minus 1 plus 0 plus 2, which is one which is a real number. Now, we do something different.

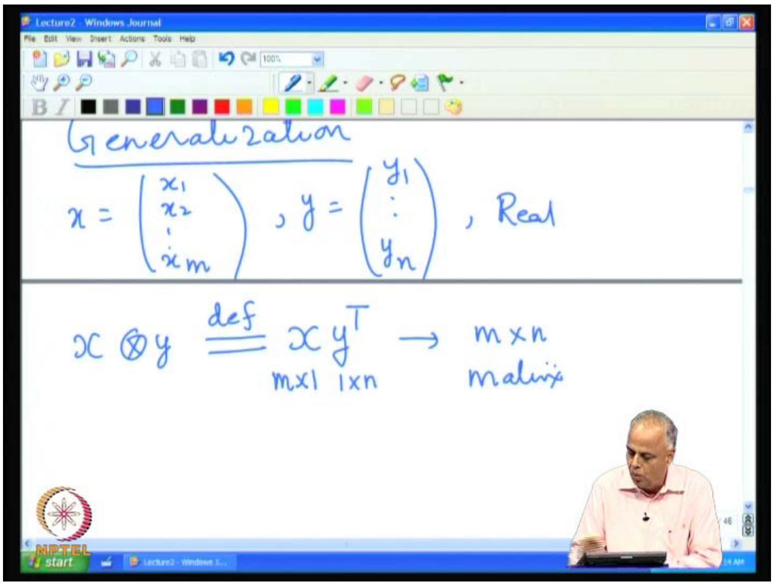
(Refer Slide Time: 13:13)



We take another type of product between two vectors and this is called the tensor product. Let us once again, start with a vector x or a column matrix x and another column matrix. Once again, let us take all are real. So, we take two column matrices x 1 x 2 x n y 1 y 2 y n both of them, having exactly n entries. Then we define, the tensor product of x with y as, we denote it by this symbol x with this into symbol enclose by a circle; that is the symbol for the dot tensor product and we define, it to be x y transpose.

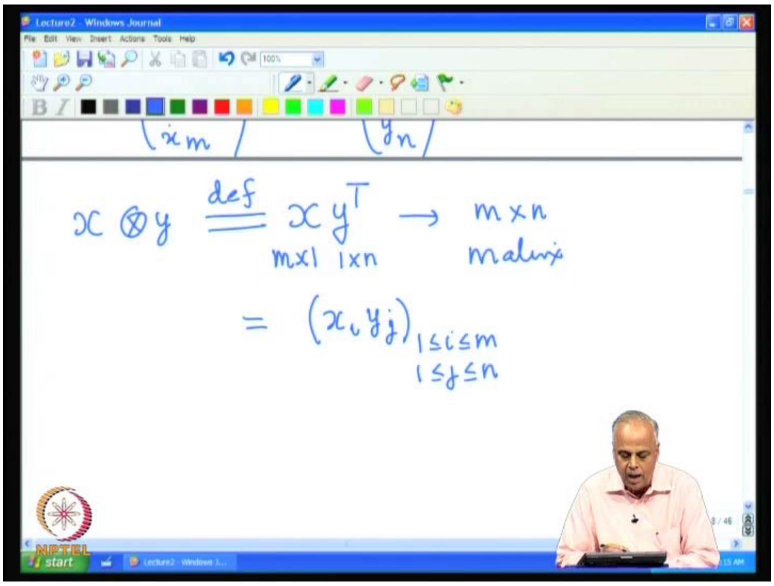
Now, since x is an n by 1 matrix and y is a n by 1 matrix, y transpose will be a 1 by n matrix. So, the product will be an n by n matrix. So now, previously when we took the inner product, the inner product up to column matrices was a number. Now, we go higher in the hierarchy, when we take the inner product or the tensor product of two such matrices, we get an m by n matrix column matrices now expand and we get, n by m matrix. Thus, the tensor product gives as a matrix. Now, supposing I can I can instead of taking an n by m matrix and an m by n matrix or taking two column vectors of the same size, we take two column vectors of different size.

(Refer Slide Time: 15:35)



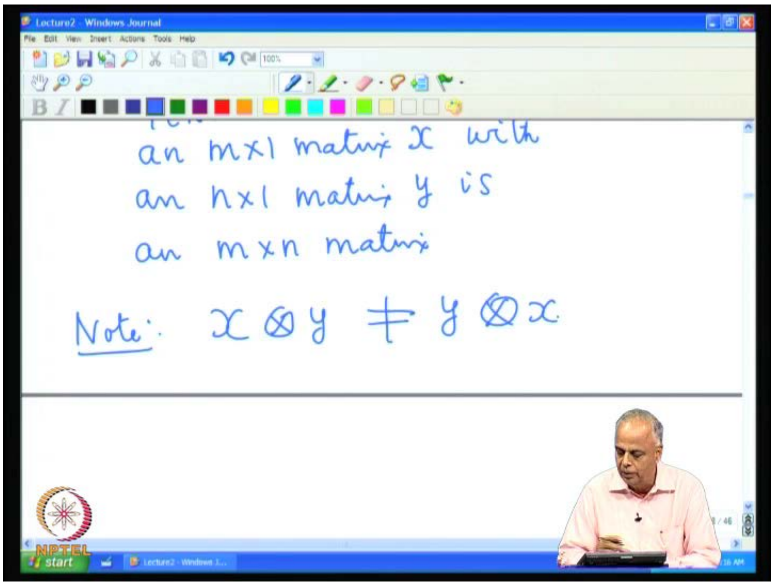
So, that is the generalization we will do, what we do now is? We take a column matrix, which is x 1 x 2 up to x m and another column matrix, which is y 1 y 2 y n again, all are real. So, one of them has m, the other one has n and now, we defined the tensor product of x with y; the tensor product of two such matrices or two such column vectors. We define, it to be x y transpose as before. Now, what do we get this term? Again x is m by 1, y is n by 1 so y transpose is 1 by n. So, the product is therefore, an m by n matrix.

(Refer Slide Time: 16:49)



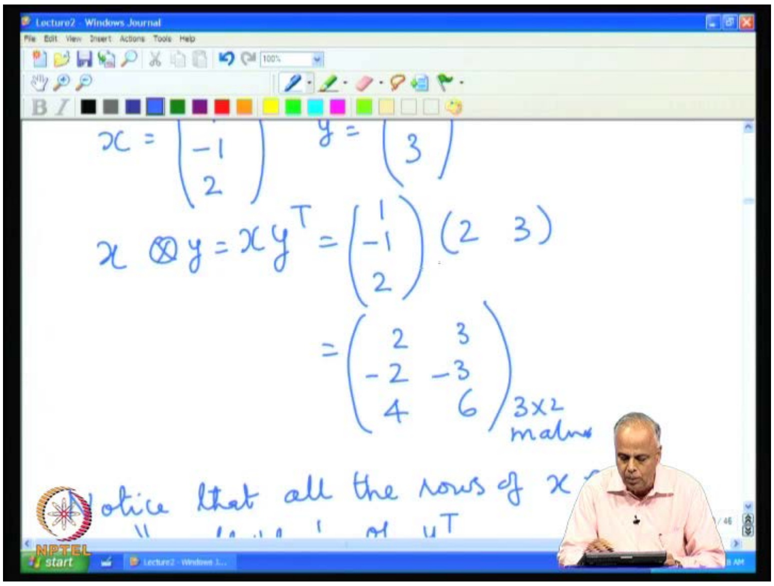
So, what is our conclusion? What are the entries of this matrix? This is a matrix, whose i j th entry is x i y j this is a matrix, whose entry in the i th row and the j th column is precisely the product of x i and y j.

(Refer Slide Time: 17:17)



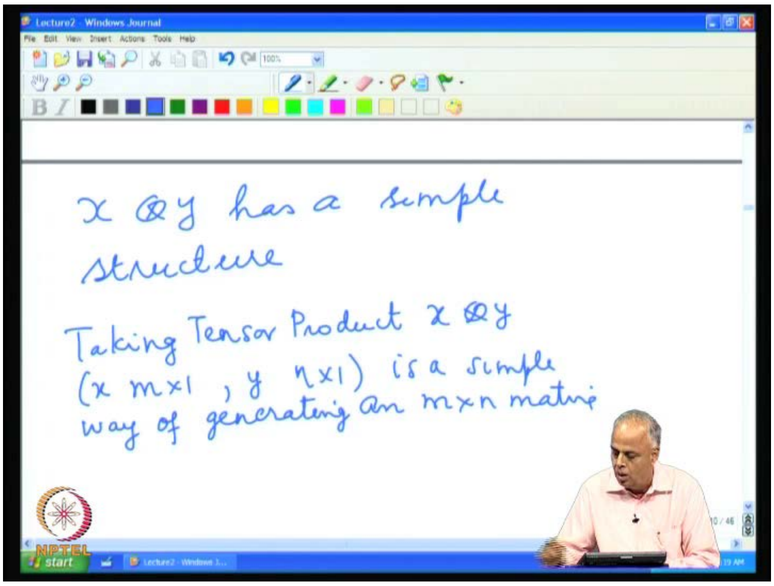
Thus, the conclusion is the tensor product of an m by 1 matrix x with an n by 1 matrix y is an m by n matrix. Note, the tensor product of x with y is not necessarily equal to the tensor product of y with x is non commutative product between matrices and which generates a higher dimensional matrix.

(Refer Slide Time: 18:22)



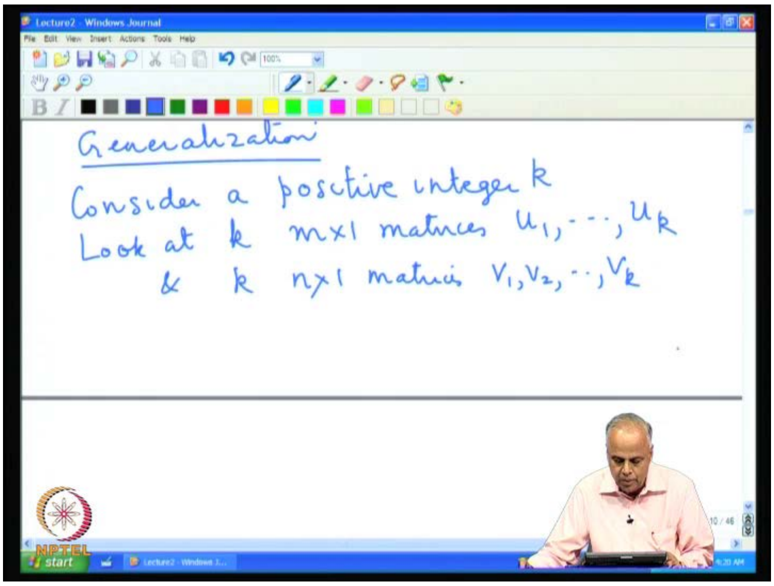
Let us look at a simple example, let s take, x to be equal to 1 minus 1 2 and y to be 2 3 then, what is the tensile product of x with y? You must take x y transpose which is 1 minus 1 2 2 3 and when we take the product, we get 2 3 minus 2 3 4 6 which is now, a 3 by 2 matrix. Notice that, all the rows of the tensor product x y are multiples of y transpose. For example, this row this row here give the just 1 multiple of 2 3; this row here is the minus 1 multiple of 2 3; this row here is the double twice multiple of this 2 3.

(Refer Slide Time: 20:00)



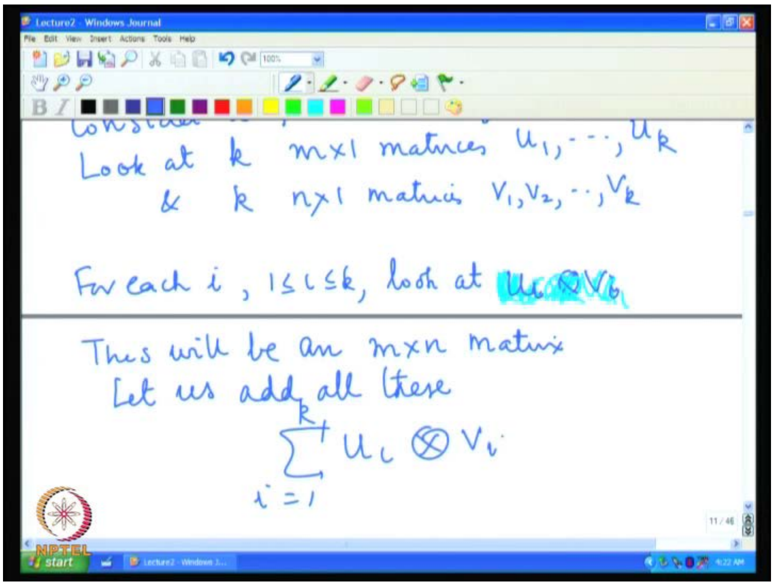
So, thus is the x x the tensor product x y has a simple structure. Namely, all his rows are multiples of just y transpose; this is one simple way of generating an m by n matrix. Taking tensor product x y, where x is m by 1 and y is n by 1 is a simple way of generating an m by n matrix and at the matrix that, we are generated as a very simple structure as we absorb, all the rows are just multiples of the row the row vector y transpose or the row matrix y transpose.

(Refer Slide Time: 21:28)



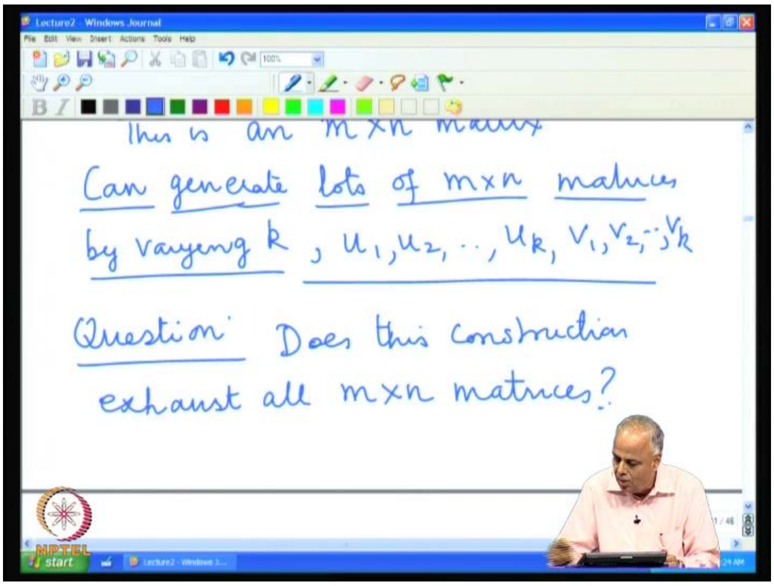
Now, let us push this construction little further. So, further generalization, instead of taking 1 x vector of m by 1 size, a y vector of n by 1 size. Let us, take a number of them and put them together, what do we mean by this? So, consider the positive integer k, look at k m by 1 matrices or column vectors u 1 u 2 u k and k n by 1 matrices v 1 v 2 v k. Now, we will form the tensor product of u with v 1 u 2 with v 2 u k with v k.

(Refer Slide Time: 22:34)



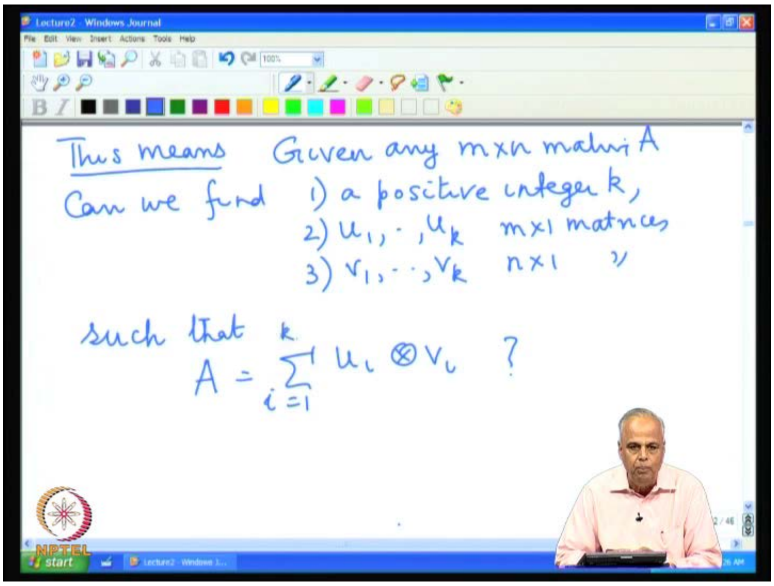
So, for each i 1 less than equal to i less than equal to k, look at the tensor product of x i with y i. Now since, x i is m by 1 y i is n by 1, the tensor product will be an m by n matrix, this will be an m by n matrix. Now, for i equal to 1, we get 1 m by n matrix for equal to 2; we get 1 m by n matrix and so on. Finally, for i equal to k, we get 1 m by n matrix. If you add all this, we still get an m by n matrix, i am sorry this should be u i tensor product. Now, for each i we get a m by n matrix denoted by this tensor product, now let us add all these we get, summation i equal to 1 to k u i tensor product with V i.

(Refer Slide Time: 24:02)



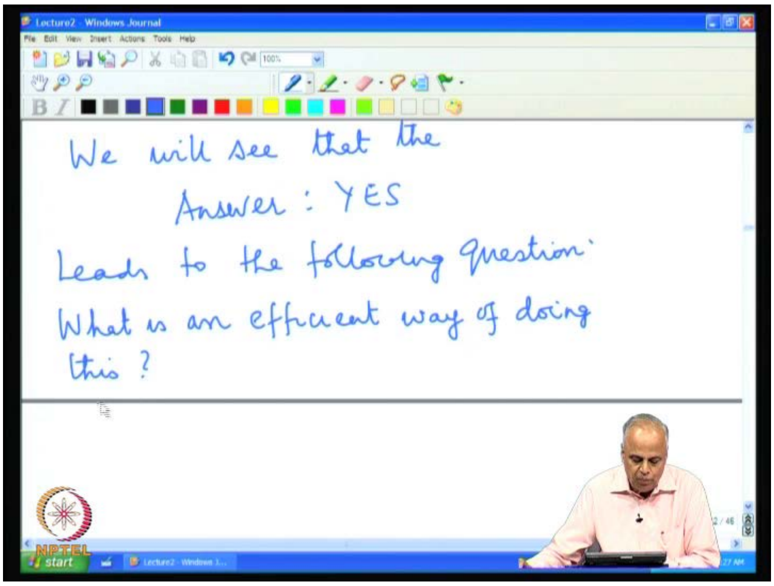
Since, each term is an m by n matrix; this is an m by n matrix. Now, we see that, we are able to generate m by n matrices by such sums of product; we can get, we can generate lots of m by n matrices by varying the number k, that is whether you would like to add k such tensor product; 1 such tensor product 2 3 4 and so on. So, you can choose the number of tensor products that you use, that can be changed and we can change the u 1, u 2, u k; we can vary the v 1, v 2, v k. So, we have lots of option of changing these combinations and each time, we may end up generating more and more different m by n matrices. So, we can generate a lot of m by n matrices by this process of tensor product sum. The fundamental question is, does this construction exhaust all m by n matrices does this construction exhaust all m by n matrices? What does this mean? This means, given any m by n matrix A.

(Refer Slide Time: 25:55)



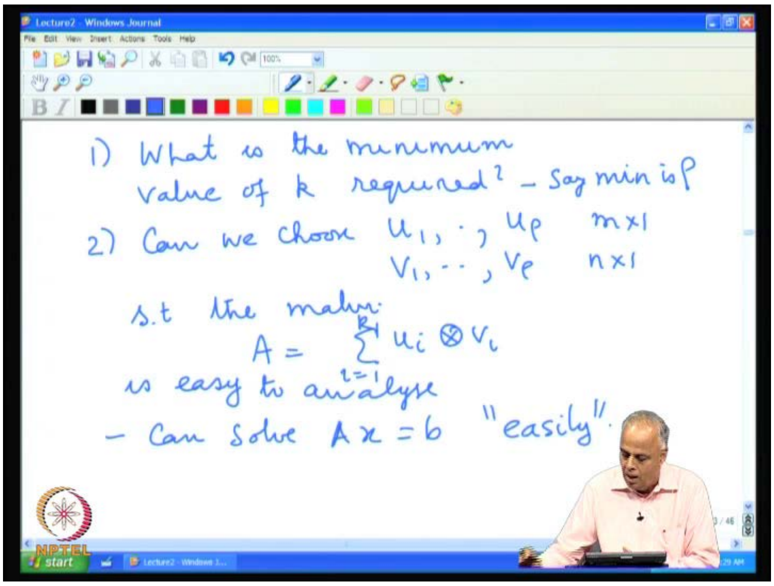
Can we find one, a positive integer k; two column matrices u 1, u 2, u k, m by 1 matrices, that is m by 1 column matrices; v 1, v 2, v k n by 1 column matrices such that; A can be now expressed as, the sum of all this tensor products u i tensor product v i. This is the question. Now, if we can do it? We can see that, A has been split into a number of small simple matrices because each one of these tensor products has a simple structure namely, every row is a multiple of a fixed row; namely, the i th row is a multiple of v i transpose and therefore, that would break the matrix into simpler parts.

(Refer Slide Time: 27:32)



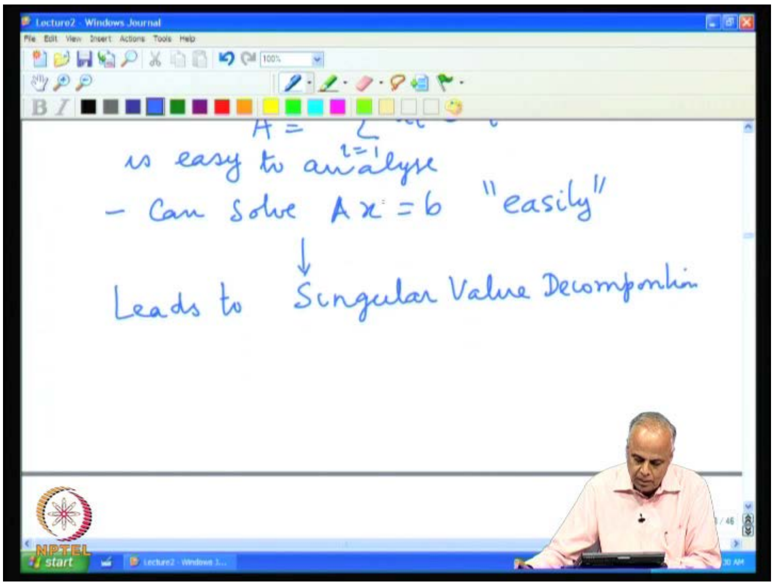
The answer we will see that, the answer is yes; this raises leads to the following question. Now that, we are assure that the matrix A, an m by n matrix A can be broken into such small pieces. Each one of which is simple, you would like to do it efficient way. So, what is an efficient way of doing this? That is, what is an efficient way of expressing a given matrix A at the sum of tensor product? First of all, we would like to have as few terms as possible.

(Refer Slide Time: 28:41)



So that is, the first question is what is the minimum value of k required? That is, what is a minimum number of sums that you have to take in order to construct the matrix A? Of course, the answer will dependent on, what is the matrix? The properties of that matrix so what is this number k will depend on A? The second question is having fixed that number k, say minimum is row. Usual eventually, it is see that, row is the rank of the matrix. Now, having fix that minimum number, can we choose u 1, u 2, u rho, these are the m by 1 matrices v 1, v 2, v rho n by 1 matrix that, will appear in the tensor product. Such that, the matrix A equal to i equal to 1 to k u i v i is now, easy to analyze. What do you mean by easy to analyze? For example, can solve A x equal to b using this decomposition, expressing A as the sum of the tensor product, we can solve a equal to b easily.

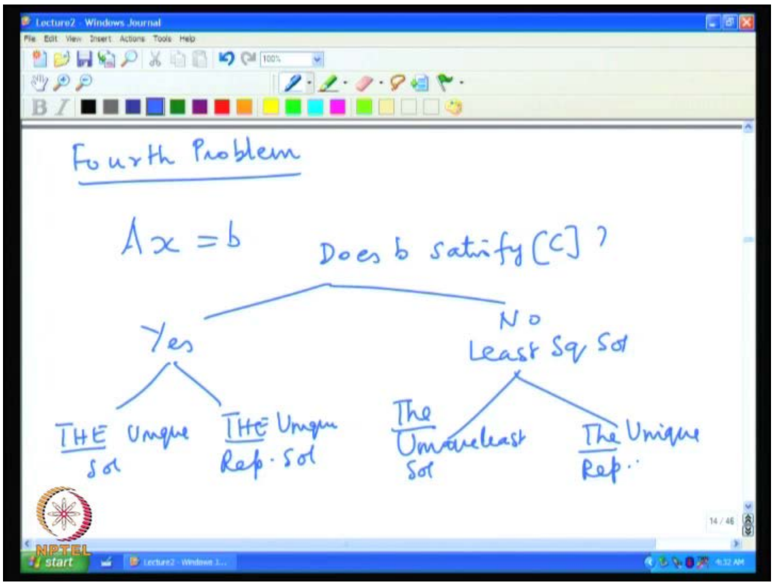
(Refer Slide Time: 30:31)



Now, an answer to this question will again, leads to what is known as the singular value decomposition? We have already seen, a singular value decomposition earlier, there we were trying to express the matrix as the product of matrices. Here, we are trying to express the matrix A as the sum of matrix, both have closely related to each other. We will when we get to these details; we will see all these are totally related to each other. This is another important problem in particular, the way we choose the u i and the v i would make not only solve in the system ax to be easy to solve, it would even help us in compressing that the data A stores. Before we begin, over formal study of Linear

Algebra, we would look at one more important problem connected with matrices or one more important problem of linear algebra.

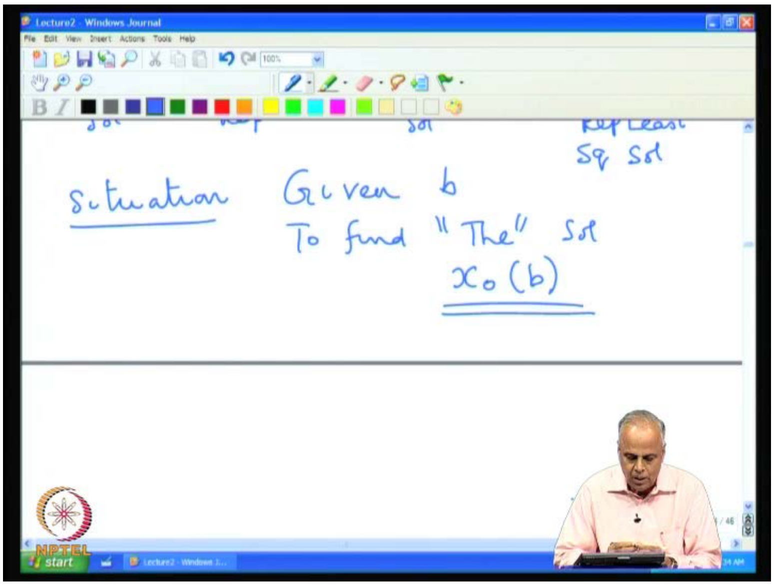
(Refer Slide Time: 31:39)



This will call, at the fourth problem. This is again, connected directly to the systems of equations and will later, will find its connection to the singular value decomposition etcetera. Let us recall that, when we have a system A x equal to b; we had situations like does b satisfy the condition C for existence of a solution; we had two situations yes and no and the yes led to two situations, the unique solution and the unique representative solution. That is in the case, when there are infinite number of solution we had to pick one of them by using some criterion. So, that our final answer is no vector x.

In the case of no, we had least square solution and this again gave us two situations, unique least square solution, the unique least square solution and the unique representative unique representative least square solution. By we see that, finally, in all the situation, we are looking at the unique solution. In this case, the first case, the unique representative solution, the unique least square solution, the unique representative least square solution therefore, in all this finally, we want the answer to be one unique vector.

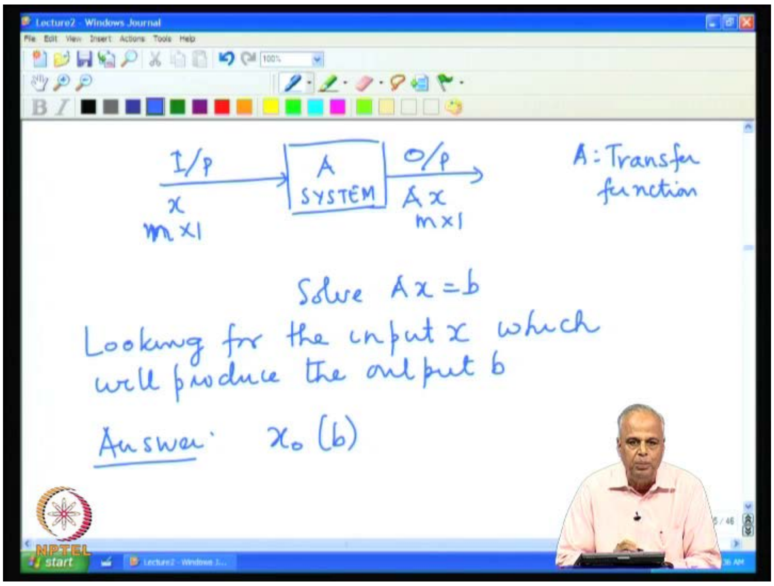
(Refer Slide Time: 31:51)



So, we have this situation. The situation is we are given b, we want to find in each one of this case the solution. When we say, the solution; we mean, the unique solution in the first case; the unique representative solution in the second case; the unique the least square solution the third case and the unique representative solution in the fourth case.

You will in general, denote this by x naught (b) because it will depend on b and it is a vector or a solution matrix m by 1 matrix m by 1 matrix and will denote it by x naught (b). So, our answer is finally, to find x naught b. This x naught b again, a reprove will represent the unique solution whenever, the unique solution exists; the unique representative solution whenever, there are in infinite number solutions; the unique least square solution when the least square solution exists only an unique or the unique representative least square solution then, there are an infinite number of least square solution and we pick, the unique representative. So, our final job is we are given the vector b produce b the vector x 0 (b), which is the ultimate answer to have question about solving the system of equation. We can view this problem in the following way.

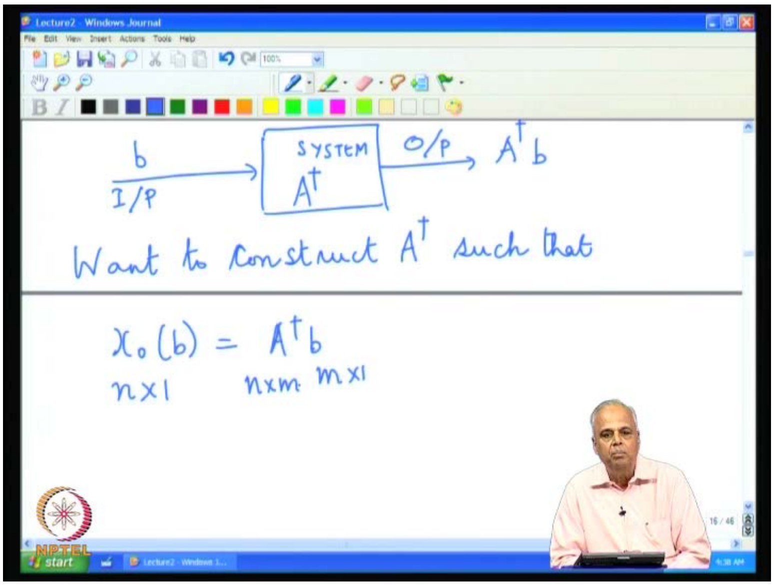
(Refer Slide Time: 35:38)



Namely, let us look at the matrix as, the transfer matrix or the transfer function of a linear system. What does it do? When we have an input x, which is an n by 1 vector, so the permissible inputs for the system are all n by 1 matrices, then the input and m by 1 matrix; the output is going to be A x, which is an input. If you put input n by 1, the output is going to be m by 1. What we are asking is, when we solve A x equal to b means, we are asking for a particular output b, and we would like to know what the input x is. So, we are looking for the input x, which will produce the output b. Our answer should be that x naught (b). What do you mean by x naught (b)? Whenever there is only one input that, produced the desired output b, x naught (b) will be that; whenever there are several inputs which produce the same output b, then the x naught (b) will be the representative input, which produces b; whenever there is no input x, which produces the output b, then we will be looking for inputs, which will take you as close to the required output b, that is that minimizes the error.

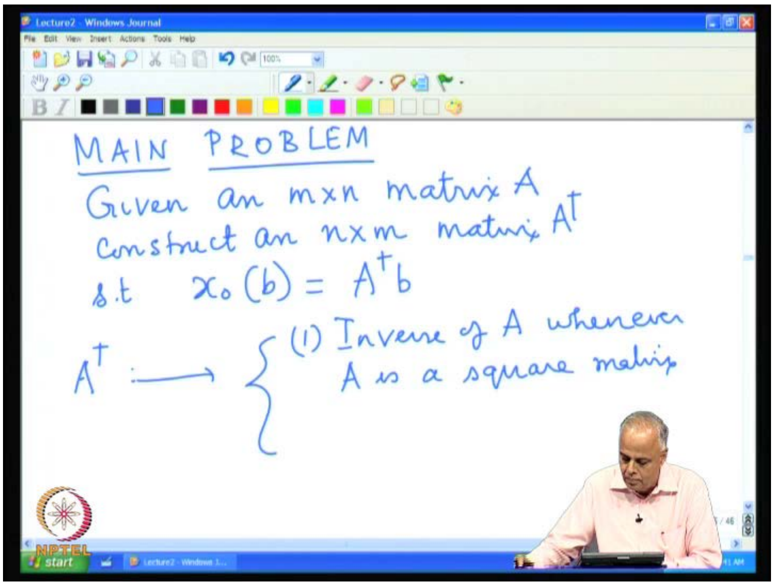
So, whenever there is a unique minimize or the unique least square solution, that input will be the x naught (b). Whenever there are several inputs which give the same minimal error then, x naught (b) will be the unique representative input. So, x naught (b) is our final answer; if you want the output b, x naught (b) is the best that can produce it. It can probably go to b and nobody else can give the better answer than that, that is what the final answer should be so. We are given b from that we are trying to construct the x naught (b). So, we are given the required output b, you would like to construct the required input x naught (b). So, in other words we are trying to control the output b.

(Refer Slide Time: 38:43)



So, therefore, what you would like is, we are only given b; we had to do some calculations on this d to produce the x naught (b). So, this boils down to, we are constructing another system for which, that another transfer function. We do not know what it is? We denote it by A dagger for which, we are going to put the input b and there is going to be an output because of this which is A dagger b. We want to construct A dagger such that, this A dagger b is our ultimate answer to our linear system. Naturally, we are expecting x 0 to be n by 1, b is m by 1. So, therefore, A dagger has to be constructed as an n by m matrix. So, the fundamental problem therefore here, the construction about the system of equations is the construction of this matrix A.

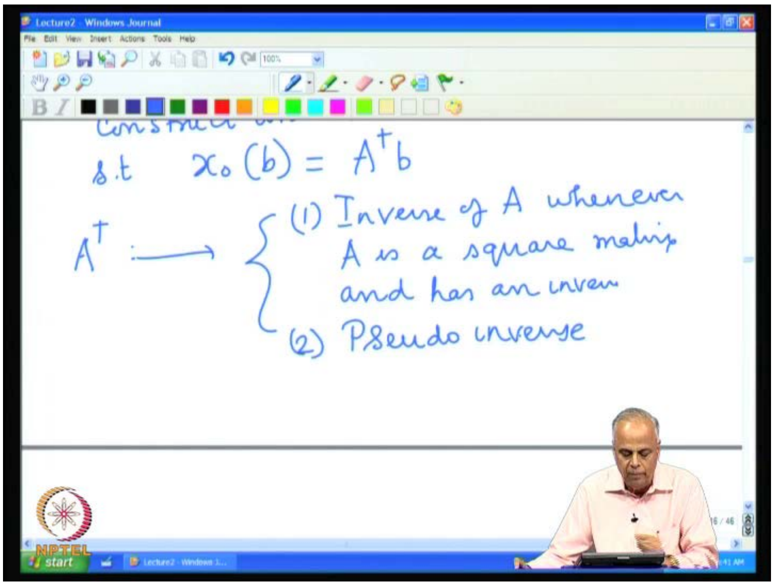
(Refer Slide Time: 40:08)



So, main problem is given, an m by n matrix A. Construct an n by m matrix, A dagger such that, x naught (b) is equal to A dagger b. I repeat again, x naught (b) as different commutations and different situation. When there is unique solution for A x equal to b, x naught (b) represents the unique solution; when there are infinite number solutions for A x equal to b x naught (b) represents the unique representative solution; when there are more solutions for A x equal to b and if there is only a unique least square solution x naught (b) represents the unique least square solution and when the only least square solution, but infinite number of them then x naught (b) represent the unique representative least square solution.

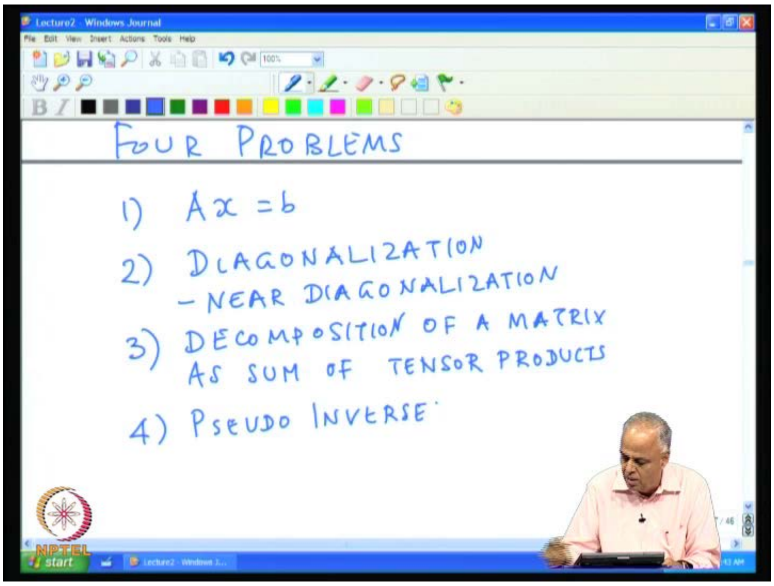
So, the fundamental problem therefore is, in the system; in the given the matrix A which is m by n, construct the matrix A dagger which is n by m. So, that x naught (b) is equal to A dagger. So, the moment pray you are b, I will send it through a dagger and produce my answer x 0 b. Which is the best answer possible for the system A x equal to b, this A dagger is reaches to several things one, inverse of A whenever A is a square matrix.

(Refer Slide Time: 40:08)



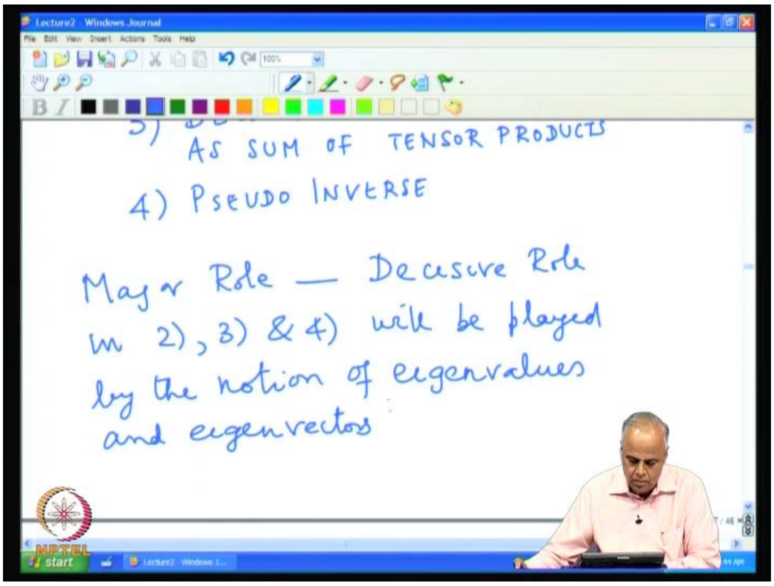
And has an inverse and in the general case, whether it is an inverse; whether it is a square; whether it is rectangle etcetera, it is called the pseudo inverse. So, the pseudo inverse into same as the inverse, whenever quit makes sense. So, this leads us to the notion of the pseudo inverse of a matrix.

(Refer Slide Time: 42:53)



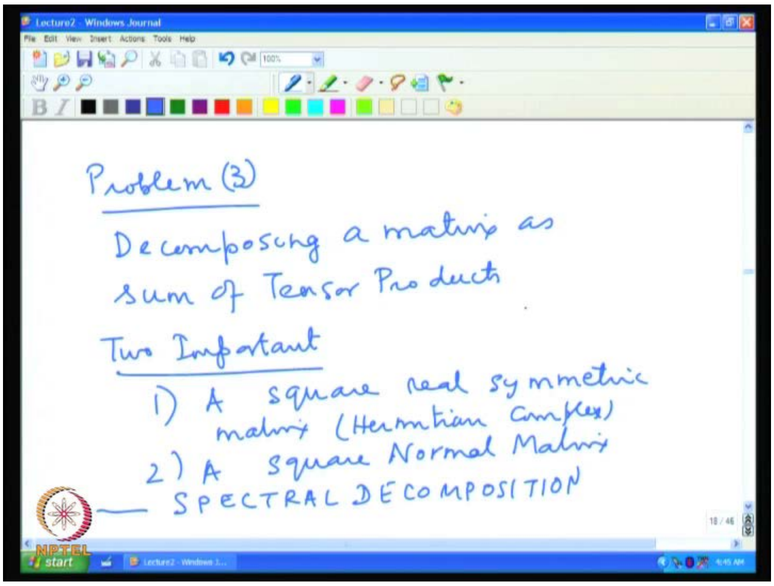
Thus, we have look at four problems one, the system A x equal to b the four problem we have studied are to generally discussed. We shall going to form, the back bone of the various studies that we have going to do. Then, we looked at the question of diagonalization and near diagonalization; then, we look at the question of decomposition of a matrix as sum of tensor products and finally, the notion of the pseudo inverse. Now, we should not loose site of the fact that, the second third and fourth problems are confidential the question that, we ask in the first problem for the linear systems of equation. So, there are all highly inter related problems and go hand in hand, one helping the other or one leading to the other. in the

(Refer Slide Time: 44:37)



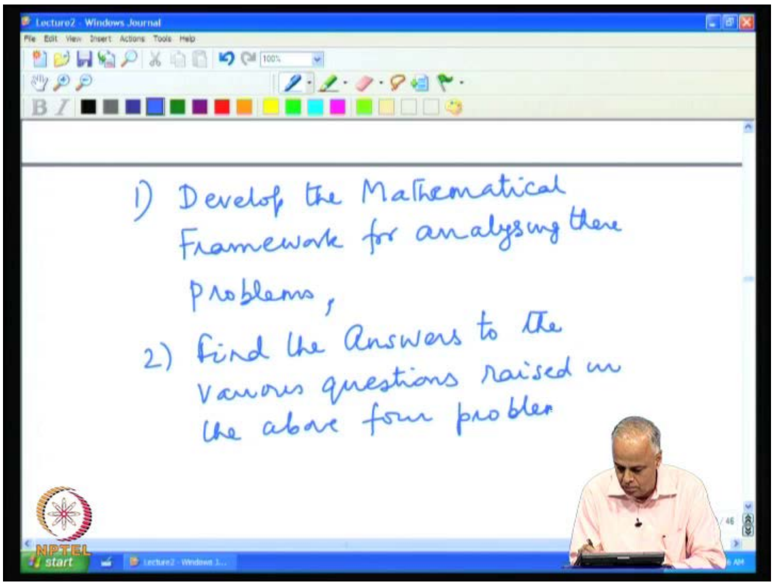
Therefore, in answering 2 3 and 4, the major role or the decisive role in 2 3 and 4 will be played by the notion of Eigen values and Eigen vectors.

(Refer Slide Time: 45:23)



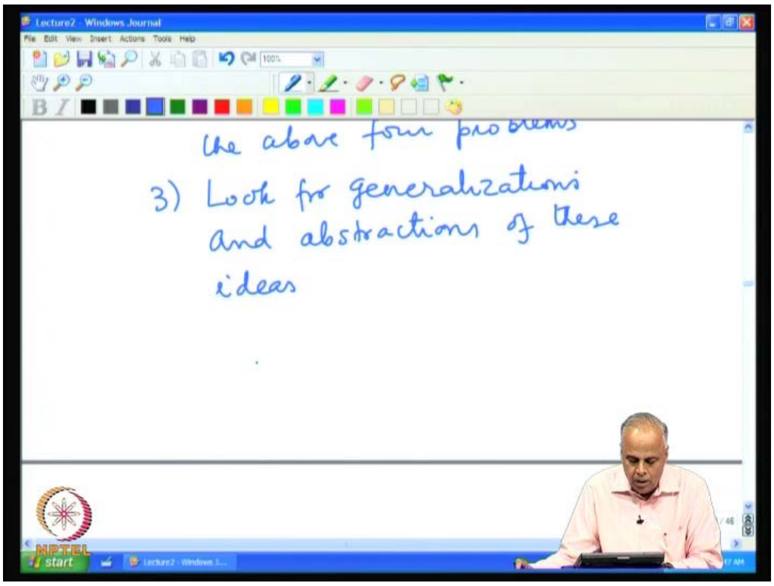
We look at the problem 3, in which, we are decomposing a matrix as sum of tensor products. Two important situations in these are the following one, when A is a square real symmetric matrix; two, A what is known as, we look at later the definition A square normal matrix. In case of complex, we should look at hermition complex matrix. This decomposition, what is known as the spectral decomposition for these matrices? As absorbed above, the Eigen values and Eigen vectors pay a crucial role in all this analysis.

(Refer Slide Time: 46:52)



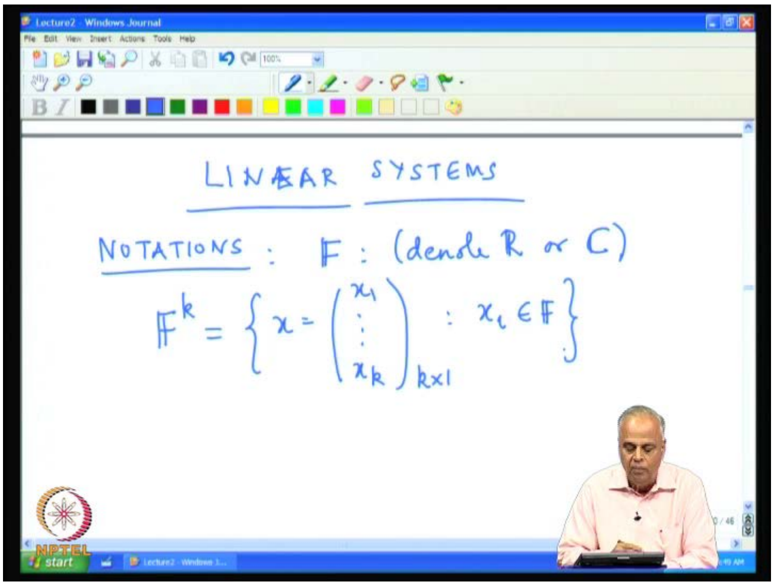
Our goal, for the course shall be the following; one, develop the mathematical frame work for analyzing these problems; two, find the answers to the various question raised in the above problems.

(Refer Slide Time: 48:02)



Three, look for generalizations and abstractions of these ideas. The main discussions in the course will be driven by these three goals. All our attention will always be too eventually to attain these three goals. Now ready to became a formal course on Linear Algebra, we have what we have seen is only a normal view of the important problems that lead as to study all these aspects of Linear Algebra that reserve the convicts on this course. So, before we begin our course, we shall give some standard notations that we will be using in this course.

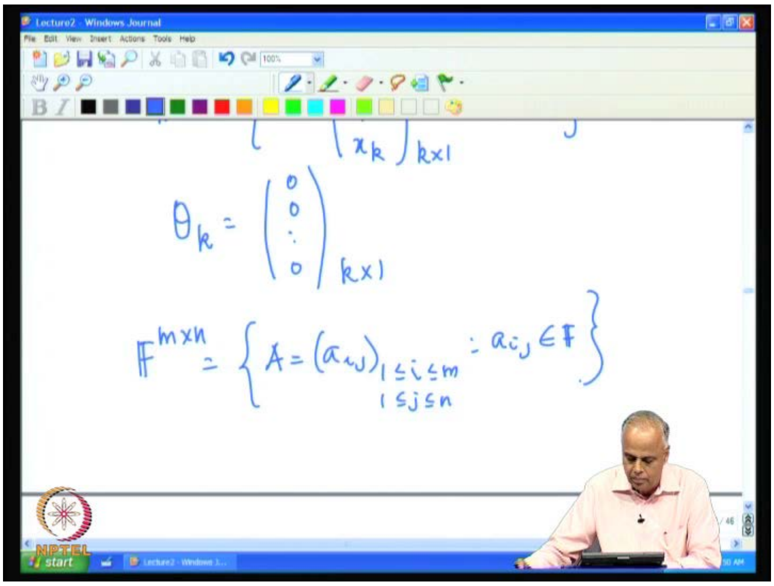
(Refer Slide Time: 49:12)



So, the first thing that we will study will be the linear systems. The notations that we will use are the following. Most of the time, we shall be dealing with matrices whose entries are real or complex numbers. So, what we will do is? We will use the general notation F, which may denote the real numbers R or the complex numbers C. R denotes the set of all real numbers, C denotes the set of all complex numbers.

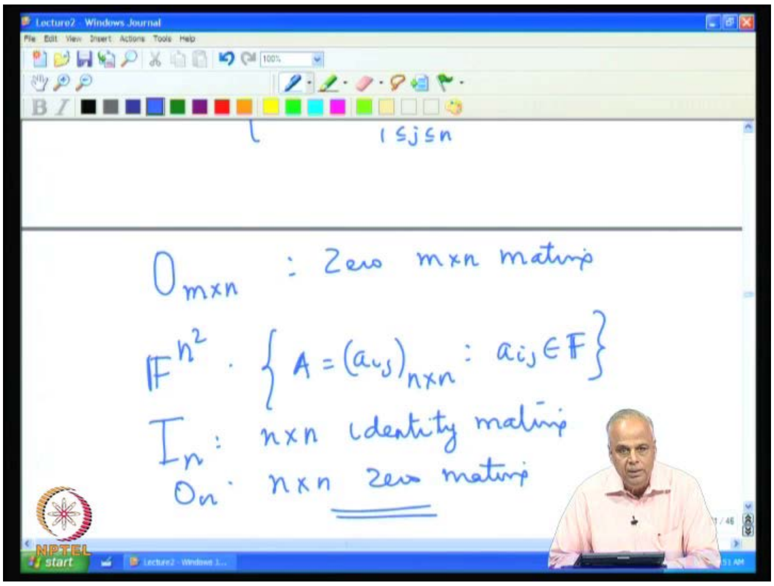
Whenever necessary, you will mention very specifically, whether we need R or C. If you do not mention then, it could be; it will work with F as R or it will also work with F as C. Then by F k, where k is the positive integer, we shall denote all column matrices with k entries that is, k by 1 column matrices where, all this entries are from F. So, when we say R k you mean all the entries are real numbers; when we say C k there is all the entries could be complex numbers.

(Refer Slide Time: 50:44)



In particular, by theta k you will denote the vector or the matrix k by 1 matrix whose, entries are all 0. Then by F m by n, we shall denote all matrices which are m by n matrices and all the entries are from F.

(Refer Slide Time: 51:21)



And in particular by 0 m by n, we meant zero m by n matrix; that is m by n matrix, all of whose entries are 0. And in particular when m equal to n, we get F n square the set of all n by m matrices such that, a i j epsilon F; and in this situation, we shall denote by I n the n by n identity matrix, 0 n to be n by n zero matrix. With this notation, we shall begin our next lecture, the study of linear system and equations.