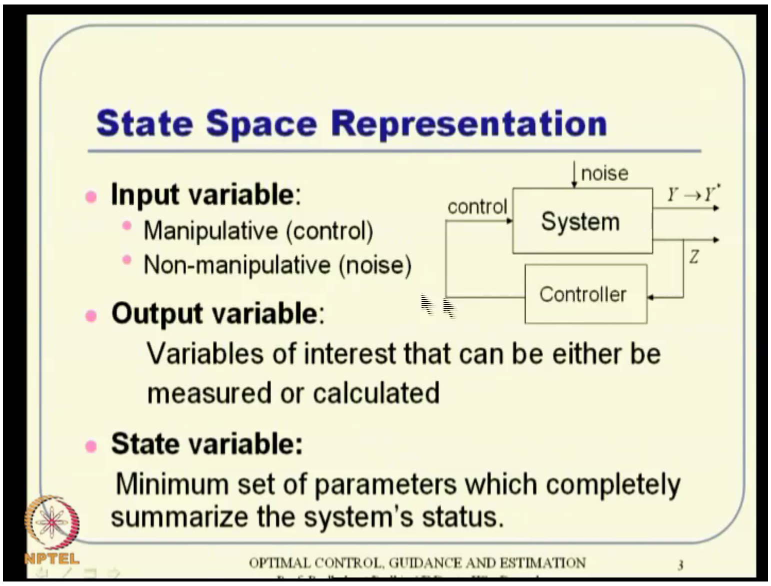
Optimal Control, Guidance and Estimation
Prof. Radhakant Padhi
Department of Aerospace Engineering
Indian Institute of Science, Bangalore

Lecture No. # 02
Overview of SS Approach and Matrix Theory

Hello everybody; we will continue with our lectures on optimal guidance, I mean control guidance and estimation course, we are in lecture number 2. The very first class, we saw some, some, overview and then some motivations of this course and all, all. This particular lecture will have some overview of state space approach and little bit over matrix theory. So, these are the two things that we, we, used it very expensively here in this course. So, we thought we will just have an overview of that. Assume that you must have already taken some course on, on, this subject or this is just to kind of a refresh your knowledge on that.

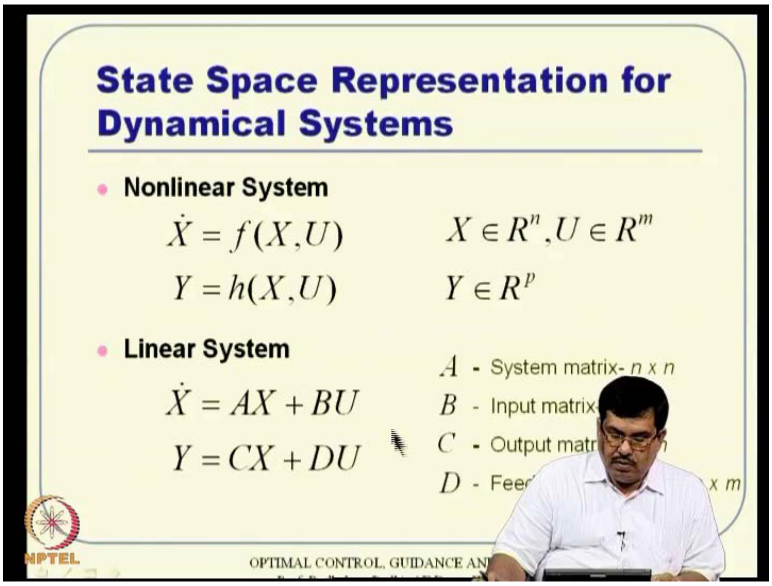
(Refer slide Time: 00:56)



So, we start with state space representative, very basic ideas and all that. So, when you talk about any system, I mean dynamical systems, we typically talk about three types of variables; one is input variable and the other one is output variable, the third one is state variable, and I kind of discuss that in the first lectures was also, and typically the input variable can be classified into two classes; one is manipulative, which is control variable or either it or it can be non manipulative which is noise, and remember, both are input to the system any way. And then, considering output variable, that can be defined as either the variables of interest or that can be like some measured variable or calculated variable. So, these two may or may not be same. This is what the diagram it tells. See in other words, your performance output can be y and that you want to drive it to some desired value Y star about the, where as the measurement output can be z, and that may, that may go into the controller and then controller is can take place actually. So, it may be so happening that part of the y variables or part or full of the, I mean all of the y variable can be part of z and vice versa probably, but they need not be same actually.

So, all these things you can be talked about in a standard state space class and all that, but anyways coming back, there is also some class, I mean this variables of great importance is state variable, and that is defined as minimum set of parameters which completely summarize this system status, and more definitions, more implications and think like that can be found in a standard text book from on systems theory, and any modern control book will kind of give you more insides with the example and all actually. But we understand what it is. It is a, what it what we mean by state variable, and that means, as a minimum number of variables, so which can describe the system state, and obviously, remember that you cannot take more than the state variable that is required, or you cannot take less than the state variable that is required either actually.

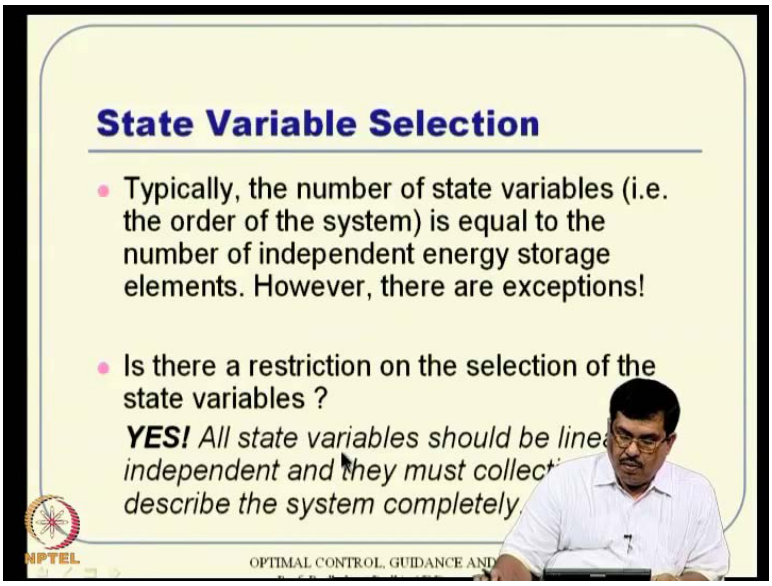
(Refer Slide Time: 03:03)



So, that is, those things are kind of non (( )). Now continuing further, we have a broadly two classes of system that we discuss about here, and one is non-linear systems and other one is linear system actually. When we talk about non-linear system, we represent the state equations this way, X dot equal f of XU. Typically time variable t is not their, where X dot represents derivative with respect to time, and remember, state variable dimensional state is R N; that means N number of states are there. Control can be RM; control belongs to RM; that means U can be of m dimensional and the output that you are, that of interest can be of P dimensional actually. Now, also you, this R, letter R represent serial numbers, so we are typically deal with all real number here. And but in non-linear systems as we know, are typically it is little bit difficult to handle. So, typical to analyze synthesize and all that.

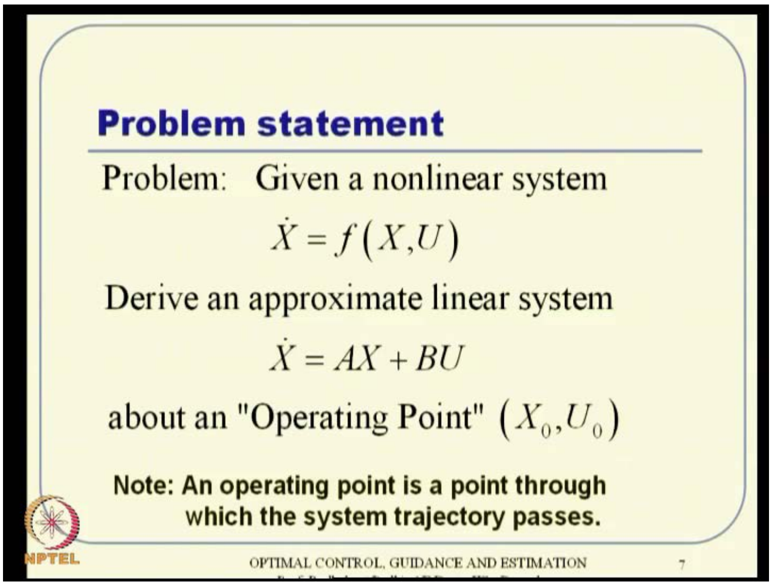
So, most of the time, you way linearize the system about some operating point and all that, and then land up with some representations like that, where X dot equals a X plus BU, Y equals to CX plus DU, but remember, this X what you see in non-linear system and this X what you see in linear system. The meaning wise they are different. When you talk about state of non-linear system, actually the two states, whereas the state of linearize system. Most of the linear systems are anyway linearize. Linearize state and controls are typically deviation variables actually; that means X represent delta X and U represents delta U implicitly here. And delta X can be the deviation of the state, from its separating point and delta U can be the deviation of control from its operating control and all that actually, so those kind of definitions are available. So, in this particular course, we will deal both with linear system as well as non-linear system, I mean we cannot, we are not talking about only one class of thing; deal with both of the things together actually.

(Refer Slide Time: 04:56)



And coming to the state variable selection, typically the number of state variable, that is the order of the system is equal to the number of independent energy storage elements, but there are exceptions as well. You can see couple of examples, in probably in some standard text books like and all that will give you that actually. But is there any restriction on the selection of the state variables? I mean the answer is yes, and the very primary requirement is, all the state variables are to be linear independent, and then all of them must collectively describe the system completely, that is, these are the two things that state variable should have actually. Now, coming to the linearization of non-linear systems, see you remember that, that is what I told you here; the linear systems are typically linearization of the non-linear systems about some operating point actually. So, how is it done typically, I mean how is it done? Let us see that.

(Refer Slide Time: 05:53)

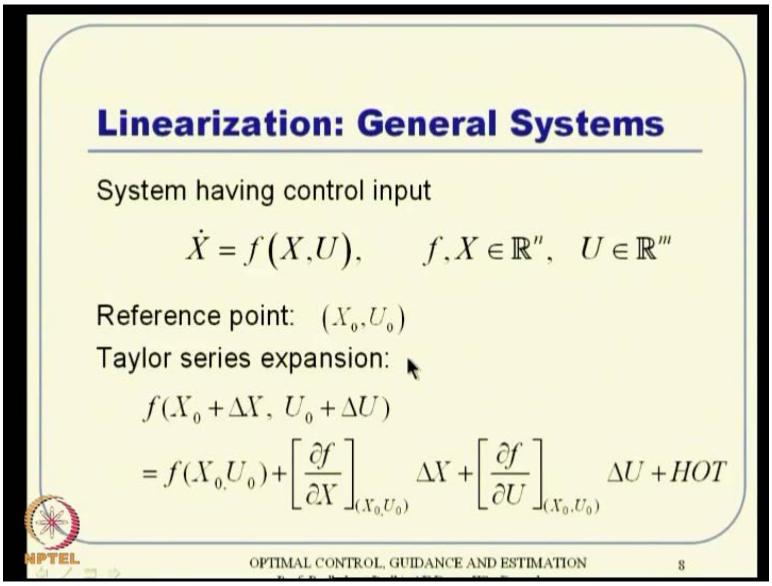


The problem statement here is given a non-linear system like this; X dot equal f of XU. Can you derive an approximate linear system about in operating point X0 U0, the approximate linear system can be described something like this actually. And also remember, when you talk about operating point, typically the miss number is, it is an equilibrium point, but it need not be true actually. So, operating point, by definition is a point through which the system trajectory passes; that means even it is passing through a tangent trajectory and all that actually, that can be a potential operating point, and so that is how sometimes we will linearize the system about a trajectory; continuously we keep on getting a new, new, new, linear, I mean linearize systems; that means the A and B matrix become time bearing and all actually, if you do that.

But if it is a equilibrium point, that is your operating point and then you linearize the system, then you land up with and be constant matrix, and this constant matrices that is why you will get into this also called linear time in variant system and all that actually. But remember that the operating point is a point, through which the system trajectory passes, that is the definition actually. And obviously, if equilibrium point is also a point through which systems trajectory passes, because if you just table system, ultimately the trajectory will come and sit there, so that is also a possibility of an operating point actually. So anyway, these are the ideas there. The operating point we noticed where, even though we represent the same state, notations state in control notation X and U.

When we talk about linear systems, by definition this state is actually deviation state, and this state is actually deviation control.

(Refer Slide Time: 07:37)



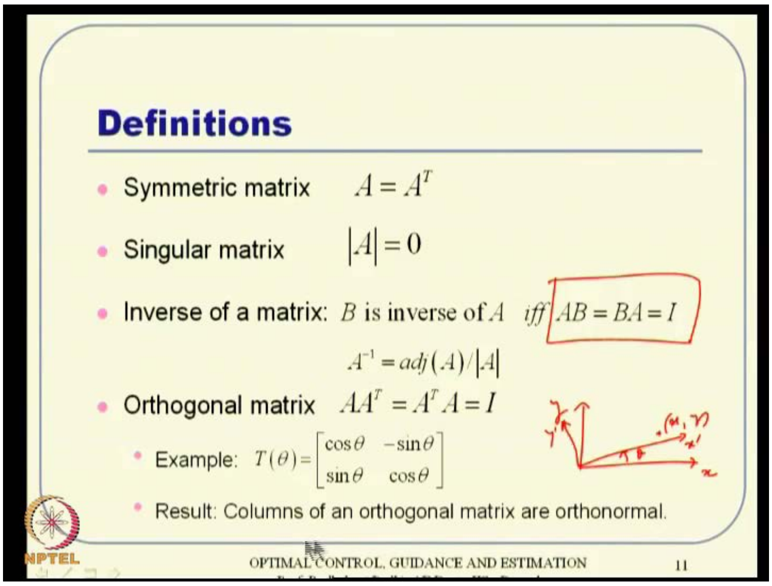
So, proceeding further, how do you get it, that is through the Taylor series expansion and all that. So here, we have a system, non-linear system, where we have a reference point here something like X not U not. So, if you put the X equal to X not plus delta X and U equal to U not plus delta U, and then, that then this side of the story f of X X U can be represented like this, and then you use the Taylor series expansion. When you can expand it something like a first term, a constant term, followed by two linear terms in, in, state in control, then this added terms. So if you, and also remember, the operating point is a point through which the system trajectory passes; that means this X not U not satisfy the differential equation; that means, if I substitute X 0 dot, X not dot, then that is nothing but f of X not U not. So, this is what I exploited here. Remember, this side of the story is like this. So, about that side of the story will be something like that, X dot equal to X 0 dot plus delta X dot.

(Refer Slide Time: 08:40)

So, that is nothing but equal to this right hand side, neglecting the higher order terms, that is approximately equal, and then we take, this two will cancel out, because this is what the operating point is all about. This is, this two will go, and then you land up with this X not equal to. Then you land up with this X not equal to, I mean delta X dot equal to a times delta X plus B times delta U. But, instead of writing down all times delta X and delta U, you think that; we implicitly redefine it as delta X equal to X, and delta U equal to U is U. Then this leads to X dot equals to AX plus BX. That is how this AX and BU represent this delta X and delta U, but also I mean if X not U not happens to be 0 0, then obviously this delta X equal to X and then delta U is also equal to U basically so; they are corollaries and all that, and remember, A and B will be defined in terms of this Jacobean matrices that way. So, we have this state equation with the system model is available to you.

So, expand this partial derivatives that way, and then put it them in a matrix form and evaluated the point of X not and U not itself. That is what you will do for A and B matrix. A happens to be NY N square matrix and B will happen to be NYM non square matrix in general basically. Now, if you land up with some system equation like that and try to carry out lot of algebra and then optimal control synthesis and think like that. Obviously, we have to deal with this matrices and we. So, we have lot of requirement for matrix and algebra on the way actually, and also we will deal with, even if you deal with non-linear system, you have to deal with gradient matrices and all that, gradient vectors something like that. So, we certainly need the concept from matrix theory actually. Let us have a quick overview of matrix theory in general, and then we can from next class onwards, we will go to optimization theory or little bit overview of numerical methods as well, before you go to optimization theory and all that actually.

(Refer Slide Time: 10:47)



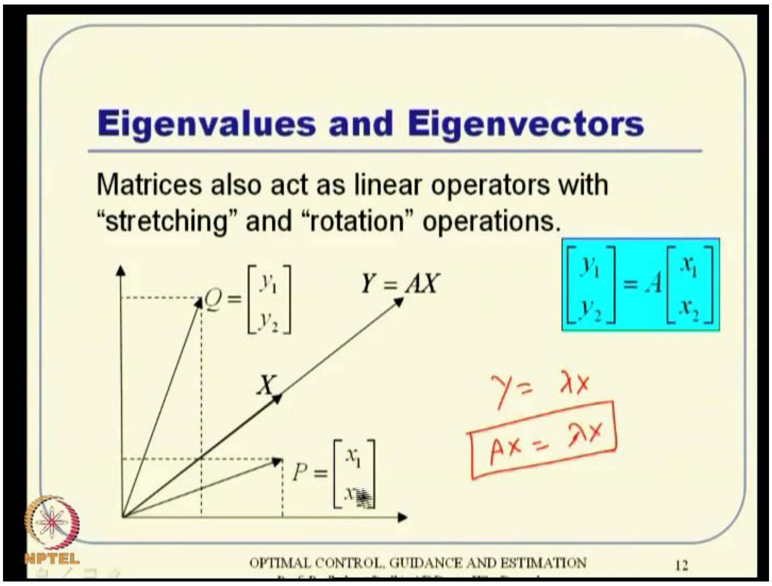
Anyway, so coming, coming to the review of matrix theory, so, let us we start with simple definitions. We all understand what is meant by matrices. So, I am not going to that level. It is a collection of numbers in an array and thing like that, but coming to some definitions, standard definition, symmetric matrix is represented as something like this. When A is equal to A transpose, then obviously A satisfies the property of symmetric. Now, when determinant of A becomes 0, and here we assume that A is a square matrix. Then we talk about well the matrix is singular matrix actually; that means inverse does not exit actually, then what is inverse. Inverse of a matrix A and B, I mean if B is an inverse of A, then this satisfies the property; A B is equal to B I is nothing but identity basically. So, that is how this forms the definition, and then, as a standard result a inverse turns out to be advent of a divided by determinant A. And again I emphasize here that do not get confused with what is definition and what is standard result actually.

As far as definition is concerned, I will put this is a definition; AB equal to BA equal to identity, and lot of this theorems, proof and then general results and all will be very easy to deal do, using this definition AB equal to BA equal to identity. However if you really want to compute an inverse for a matrix and all, probably that is a, that is a standard results that we typically use also. Even though it is not very computationally efficient way of computing A inverse by the way. So, that is the difference, so it happens to be like one of the same. In other words, both are if and only if condition sort of things. So, sometimes people get confuse that A inverse definition is this A adjacent of A divided by determinant of A, but in definition sense, this is the definition and this happens to be a standard result actually. Now, the next concept is, what is called as orthogonal matrix. An orthogonal matrix happens to be something like, orthogonal matrix happens to be something like this way. So, if A transpose equal to A transpose A equal to identity, then it is an orthogonal matrix.

And also remember that A times a transpose and A transpose A are guaranteed to be kind of, I mean square matrix, even though A is a non square, and top of that, you can also show that A transpose and A transpose both happens to be something like positive semi (()) matrix, something like that, is not that typical solve actually. Anyway, in addition to that, if they satisfy this property that A transpose equal to A transpose A equal to identity, then the matrix A is called as an orthogonal matrix. And one property of orthogonal matrix is like this; if a matrix A is orthogonal, then all its columns are suppose to be orthonormal vectors actually to each other. So, that is a vector sense basically; that means, if you have a A matrix of sum matrix which is given by some columns and all that way, and then each of the columns that you see here will be orthogonal, rather normal to each other actually in a vector sense. So, that then typical example is something like this, this is a very standard what is called as rotational matrix; that means, if you have a point, then representational turns out to be like this.

If you have something like, I mean this coordinate frame and think like that, you have this X verses Y coordinated frame, and you have a point where XY is there, and then the question here is, if the coordinate frame rotates Y angle theta let us say; that means, you have Y prime and X prime Y prime, and this is again angle theta. Then in the in the new coordinate frame X prime Y prime, what is the coordinates, I mean the values of this coordinate X prime Y prime. So then, that is kind of given by this transformation matrix and all actually; that is why it is called a rotation matrix. If you take a coordinate in the original coordinate prime and multiply with this, pre-multiply with this matrix and whatever number you get, that number happens to be your coordinate in the new coordinate frame actually. And this result I have already told you, that columns of an orthogonal matrix are nothing, but they are orthogonal to each other in a vector sense actually.

(Refer Slide Time: 15:28)



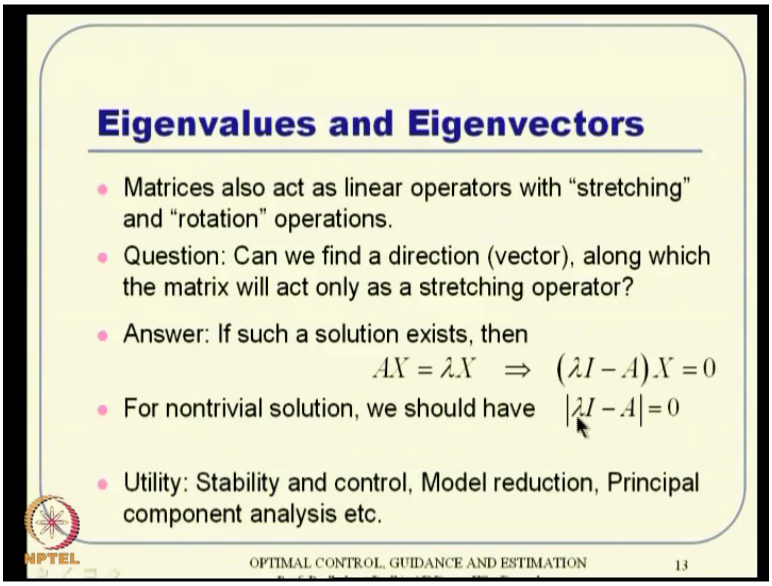
The next concept is Eigenvalues and Eigenvectors and this is how you can think of interpreting it actually. Now, this is a common mistake that we people take, if we are a eigenvector then. Well there is a standard equation for that and then I will solve it, and especially about Eigenvalues, A minus lambda is determinant equal to zero and then I will get values for lambda's and that is nothing but Eignevalues and all. So, that is the standard result again, but the concept can be understood something like this. Let us start with simple understanding here. So, let us take a point p in a coordinate frame x and y. So, and talk about this let us say this is something like x1 and x2. So, the coordinate frame happens to be something like, well whatever this is the, I mean horizontal and that has the vertical coordinate frame you can think of in general. So, we have a point p and that is x1 and x2 originally, and we multiply that by A matrix, some matrix A.

Then we will land up with some numbers y 1 and y 2, and that y 1 and y 2 can represent a point in the coordinate frame something like U. That means, if I join this origin to this point p, and origin to this point q, then I am actually getting two vectors. Now the, if you look at it slightly closely, then it turns out that, by simplify, multiplying this vector x 1 and x 2 by matrix A. What I have done. I have transformed this point p to q in general, in a way, what I am doing is, I am actually taking this vector, and then rotating it as well as stretching it. So, there is a stretching and rotation in operation involve basically. So, any matrix multiplication is, matrix are also called linear operators by the way. So, any matrix multiplication with a vector will give you another vector, and that vector need not contain the same magnitude and same direction actually.

So, it can act as something like a stretching operator as well as the rotation operators. Now, the question here is something like this. It may not happen in general, but suppose, I have a vector x some direction somewhere, and I multiply that with matrix A, I get another y which is a line along the x axis, I mean this vector, origin to x vector and it is align their; that means this matrix is operating only as a stretching vector not as a operating, not as a rotation operators, but as simply as a stretching operator. So, that means, it does not affect any other direction. It just takes the vector x and then just stretches it along further actually, and stretching by definition can means shrinking also, and so, that is also possible actually.

So, that is the question we are asking actually. So, in other words, does their exits any vector x, some vector x, for a particular matrix A, so that this operation a times x will give us vector y which is nothing but a stretched value of vector x basically, that is all we are talking. So, in other words, if you think about that, that is how this all starts actually. So, you are asking something some equation like, what are asking. If y is nothing but I can interpret that y equal to nothing but lambda x sort of thing, but y is also x remember that. So, what I am getting here is, x equal to lambda x. That is how the eigen of, I mean Eigen of, Eigen, Eigen, eigenvalue and eigenvector equation is define actually. So, y is nothing but lambda x; that means that is a stretching operator, and y by definition is a times x any way. So, that is how we get A into lambda x actually.

(Refer Slide Time: 19:37)

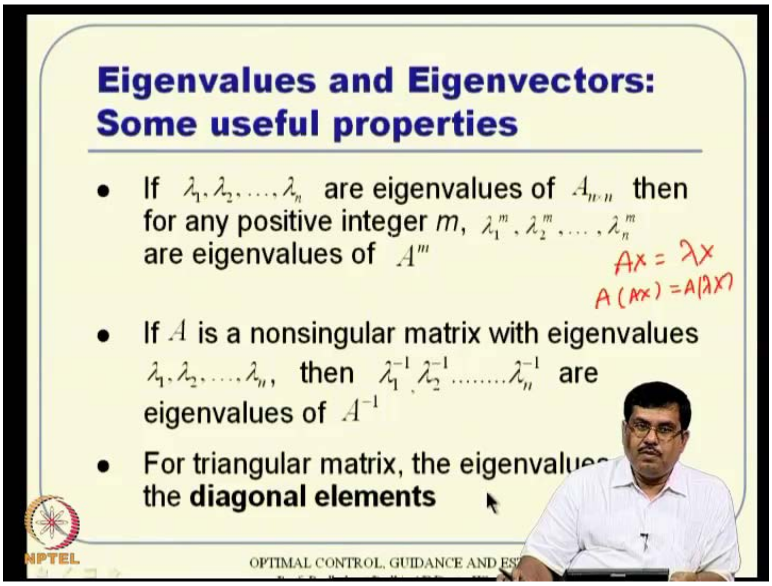


So, that is how you get it, but then the solution, the question is when does, when can we have a solution to these equation, and that solution to this equation obviously turns out that, if that is true, and that has to be true as well; lambda I minus A equal to, I mean lambda I minus A by whole multiply with x equal to zero, and if you really need to have a nontrivial solution, that means remember this equation is invariably true, if you have 0 0 0 0 as X vector; that means it is a very trivial thing, we are not moving anything away from the region actually. So, other than that, if you take any other thing, then these two have a meaningful solution and nontrivial solution. We should have this relationship satisfied. That is a standard result from linear system equations actually. That is how this equation comes here actually. So, this is, if you take this equation, remember determinant is a scalar operator.

So, you get one equation, but then, this can be the Nth order algebraic equation; that means, you can have N solutions and all that actually. So, that is how you get, if the matrix is of dimension NYN, then you land up with something like n equations, I means it is Nth order equation, and hence you will have Eigenvalues and because this coefficient will be really numbers. You will land up with either numbers or solutions or complex conservative solution. You cannot have a single complex number running out actually. So, that is again result actually. Where do we use these concepts, by the way, how do you evaluate Eigenvectors and all. Once you want to evaluate the Eigenvalues like this, you go back and put it there. Remember, it does not give you some sort of like a unique solution. It may give you a something like in other constant equation, and hence, you will have infinite solution.

Where also remember, eigenvector by definition, they do not have any magnitude; that means you are happy with only the direction of that. So that, And then, sometimes we do this normalization of the vector until ok, normalization vector and all that, but by definition, Eigenvectors do not contain or do not have any magnitude associated with them actually. Anyway, so this is how it is. Where is the utility. Utility of this analysis is huge; it is used in stability and control analysis. It is used in optimal control synthesis of linear system anyway. It is also use in model reduction principal component analysis and the variety of other applications in almost every field of engineering actually. So, that is how the importance of these concepts comes actually.

(Refer Slide Time: 22:01)



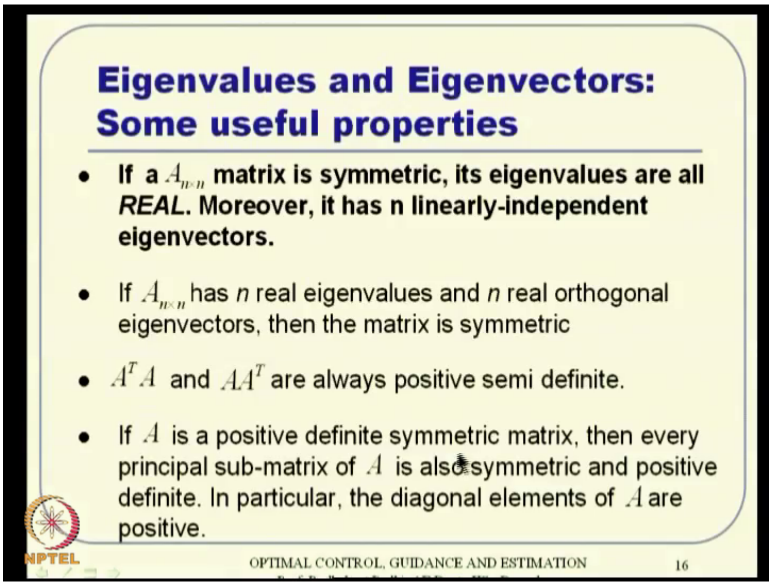
Now, there are such a very standard results that come handy also. One of the few result, I mean very standard results I will just take it through. In other words, this standard result tells you that, if you have lambda 1 to lambda n, is Eigenvalues of matrix A which is a n by n. Then for any positive integer m, you really do not have to evaluate this a to the power m and then go through this equation and all that actually. The result tells you that, if I have A to the power m, then that Eigenvalues of A to power m will satisfy this relationship. All that I have to do is just to take it, just to kind of, find of individual powers of these values and I am done actually, and it is very easy to show that, from a prove sense also, in a secondary sense I will show that. So, you start with something like lambda x equal to lambda x.

That is how lambda will satisfy Eigenvaule property of matrix A. Now, just that you do multiply both sides, that will get A into lambda x, A into. As you will get A into lambda x, and then, if you go through that and then this turns out to be something like A square x equal to lambda times Ax, and lambda times A x is again, you know this lambda times A x equal to lambda x, so that then you land up with lambda square x. So, A square x is nothing but lambda square x sort of things. So, that means it tells you the lambda square is an Eigenvalue of I square. So, that is, and then you can keep on repeating this exercise to get higher power source actually. So, let us continue, then, the next one it is very easy to show that again, that if I what about A inverse actually.

So, remember A inverse is a computational expensive operations. So, we really do not want to carry out inverse operation, unless otherwise it is a widely necessity, and also turns out that, if know the Eigenvalues of I matrix already, then the inverse matrix A inverse that will have Eigenvalues something like nothing, but lambda inverse; that means 1 by lambda sort of things. So, they are Eigenvalues of a inverse. So, again very easy to show that, by simply taking A inverse multiplying both sides by A inverse actually. We land up with x here in the left hand side, and then it will be A inverse right hand side, and then you can be very easily show, you can 1 by lambda into x equal to A inverse x sort of things here. If you do the algebra yourself and you will get convince, in a just text two line algebra actually so that.

And the third thing, third property that comes to mind is, for a triangular matrix, remember it can be either upper triangular or lower triangular, and of course, diagonal matrices are special cases of that as well, and for such matrices, the Eigenvalues are nothing but diagonal elements actually. It is extremely easy to show that again, so starting from this equation actually, if it is, this equation will not get anywhere else. You will just simply deal with this diagonal elements, and hence, you will get something like, I mean very easily you can, so that it will be multiplication of lambda minus, while the diagonal element A 1 1, lambda minus A 1 1 into lambda minus A 2 2 into lambda minus A 3 3, like that all that thing equal to zero, and hence, lambdas will be nothing but A 1 1, A 2 2, A 3 3 actually. All this happens to be diagonal elements actually.

(Refer Slide Time: 25:40)



So, this is how it can be shown this way. Now, coming to these, some of these other properties and this happens to be quite useful property actually, that if some, I mean symmetric, not symmetric. If A square matrix is symmetric, the theorem talks like this, that if a square matrix is symmetric, then its eigenvalues are all real. And remember, eigenvalues can in general be either real or complex conjugate pairs actually, but this theorem tells you that all that is necessary, for this theorem is that matrix needs to be symmetric, and if you have a symmetric matrix.

See eigenvalues are guarantee to be real actually, and it can be proved also, a little bit involved, but it can be shown with not that much of difficulty actually. And the theorem again tells you that not only that, moreover it has N linear independent eigenvector. So, these are beautiful property of a symmetric matrix. So, whenever we do not have a symmetric matrix, so we will rather try to decompose that in something like, I mean a combination of symmetric and then sque symmetric matrix and all that actually. That is a, I mean those things are standard results in linear algebra anyway, but it talks, I mean this, just notice that, I mean if you really know that a matrix is symmetric, then you can very quickly conclude, that eigenvalues are real and linear independent eigenvectors as well actually, very beautiful and very powerful property actually.

I certainly asked you not to forget this. And this next one, next theorem tells you that, if a square matrix and n real eigenvalues and n real orthogonal eigenvectors, then the matrix is symmetric as well. So, this is kind of a counter theorem for this particular thing actually. This theorem tells that if a matrix is symmetric, then this thing happens. The question is when the other thing, reverse thing happens actually, and the answer to that is here, that if a matrix has n real eigenvalues, and as well as N real orthogonal eigenvectors, then only the matrix is symmetric. So, that is property something like a converge theorem basically. Then the next one, talks about something like this, A transpose A and A transpose are always positive semi definite matrix basically. Again it is very easy to, so it is not that difficult at all.

(Refer Slide Time: 28:12)

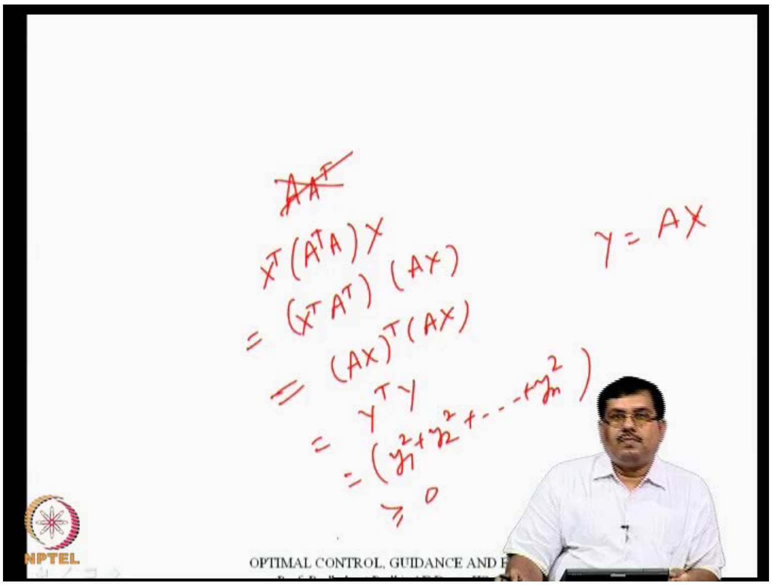
Terminology	Definition	Properties of Eigenvalues
Positive definite $A > 0$	$X^T AX > 0 \quad \forall X \neq 0$	$\lambda_i > 0, \quad \forall i$
Positive semi-definite $A \geq 0$	$X^T AX \geq 0 \quad \forall X \neq 0$	$\lambda_i \geq 0, \quad \forall i$
Negative definite $A < 0$	$X^T AX < 0 \quad \forall X \neq 0$	$\lambda_i < 0, \quad \forall i$
Negative semi-definite $A \leq 0$	$X^T AX \leq 0 \quad \forall X \neq 0$	$\lambda_i \leq 0, \quad \forall i$

NPTEL OPTIMAL CONTROL, GUIDANCE AND ESTIMATION 17

So, if you have a transpose for example, in case, what is positive semi definite and all. I think if you, it is here actually. Let us go there and then will go back to that. So, by definition, the matrix is said to be positive definite provided us, I mean it satisfies this relationship, that if I carry out this operation X transpose A X, for any nontrivial exit; that means, all that is necessary x is not on the origin actually; x is not equal to 0 0 0. I take any real number; I mean any value for X vector, and simply carry out this algebra; X transpose AX. So, this X transpose AX is remember it is one by N times N by N times N by 1; that means it is actually guarantee to a scalar quantity. So, this scalar quantity, if it is greater than zero for all such X, then the matrix is called as positive definite, and if you can tell only this much, that it can either positive are equal zero, greater than equal to zero, then it turns out to be positive semi definite.

And similarly if it is strictly negative, then it turns out negative and less than equal to zero, then it turns out to be negative definite. And again remember, this side of the thing is definition, and that these definitions are tightly related to this properties, but many times again we take this properties as the definition. When you forget that, this is where it comes actually. So, whenever there is a something to show and think like that. I certainly urge that you carry out operations along with these definitions ideas. That will be much easier. But when you want to evaluate, and if you know some values, some properties of eigenvector, I mean some values of eigenvalues and thing like that numerical values, then looking at those values you can also conclude, whether the matrix is close to definite or not all that actually. Anyway, coming back to this, the climb here what I told, is A transpose A or A transpose, both are always guaranteed to be positive semi definite at least.

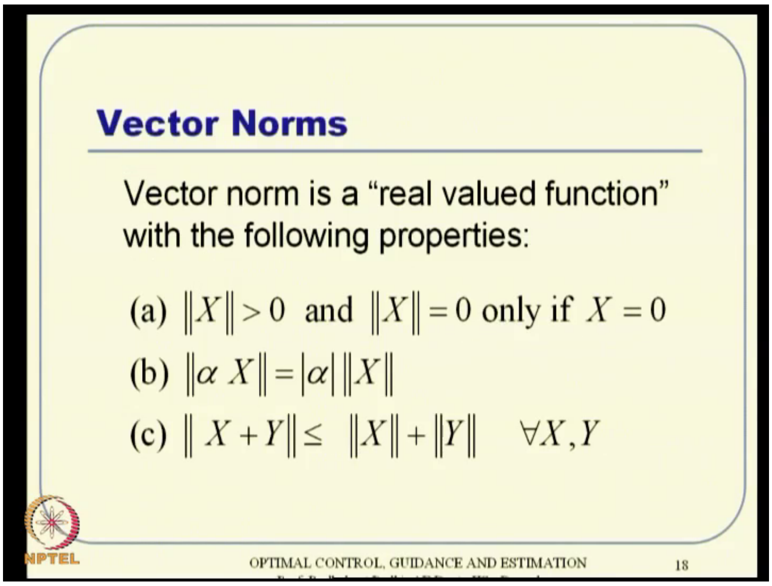
(Refer Slide Time: 30:01)



Now, it is very easy see that A transpose is guarantee to be a square matrix, no matter whatever, whether A is a square matrix or not. Now, all that you, all that both actually, I mean both at A transpose A and A transpose are guarantee to be positive semi definite. Now, all that you need to show, that suppose I want to show this as a positive semi definite. Then by definition, I have to do this operation where for all nontrivial X and that turns out to be, this is similar anyways. So, let us not vary about that. Let us carry out these one. This turns out to be X transpose A transpose into AX, and this is nothing but if take A transpose, this transpose property and all, this is nothing but AX transpose times AX, and if you define remember, AX is nothing but a vector actually; AX is a vector, so I can define that is y. So, essentially this land up with something y transpose y, and that is, if you think a little bit carefully this component sense.

It is going to be something like y 1 square plus y 2 square plus y N square (( )). So, this guarantee to be positive is a greater than or equal to zero. All these are submission of quadratic terms and things like that. So, it is very easy to show, that this A transpose A is guarantee to be positive semi definite at least actually. There is another result which tells you that, if A is a positive semi definite matrix, then every principal sub matrix of A is also symmetric and positive definite. In particular, the diagonal elements of A are positive actually. So, because we talk about this concept of sub matrix and think like that, but remember if somebody tells us that, there is a positive definite matrix say, then the diagonal elements of A are certainly going to be positive. So, that is a property that many times we use actually.

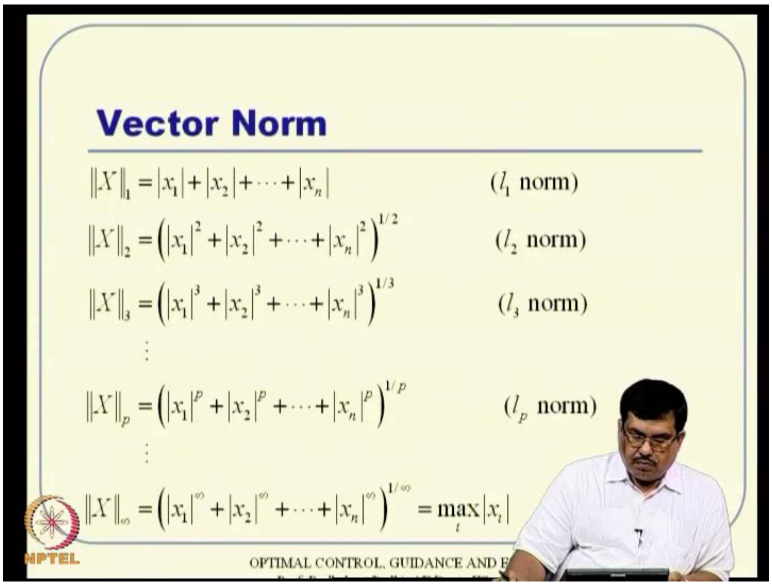
(Refer Slide Time: 32:00)



This is I have already told, when what is defined as positive definite, positive semi definite, negative definite, negative semi definite matrices actually. Then there is another concept called vector norm. Again we use it heavily in optimal control theory as well, and vector norm is defined something like that. It is a real value function; that means the output is always real number, which satisfies the following property again; that means, if I know vector x, the norm of vector x can never be, I mean never be negative, and most of the time it is positive. It is zero only if X is a trivial vector zero; that means, if only if x is 0 0 0 that origin vector, origin point rather, then only the norm is zero; otherwise is always guarantee to be positive. Essentially the norm is nothing, but the concepts of distance actually, distance on the origin something like that way.

So, distance from the origin is zero, only when you are at the origin itself. Anywhere else is the distance is certainly going to be positive, that kind of things actually. It also satisfies the second property, alpha time X. If alpha is a scalar, if you take the norm, it will satisfy modulus of alpha and times norm alpha. That means it does not matter the sign of alpha really actually, whether alpha is positive or negative, it is guaranteed to stretch it actually, and then there is also triangle inequality result which tells you that, if I take vector X and vector Y, and I sum it up and then take the norm, that is going to satisfy this, less than equal to norm X plus norm Y actually. So, that is the concept of vector norms, a meaning of having a distance actually.

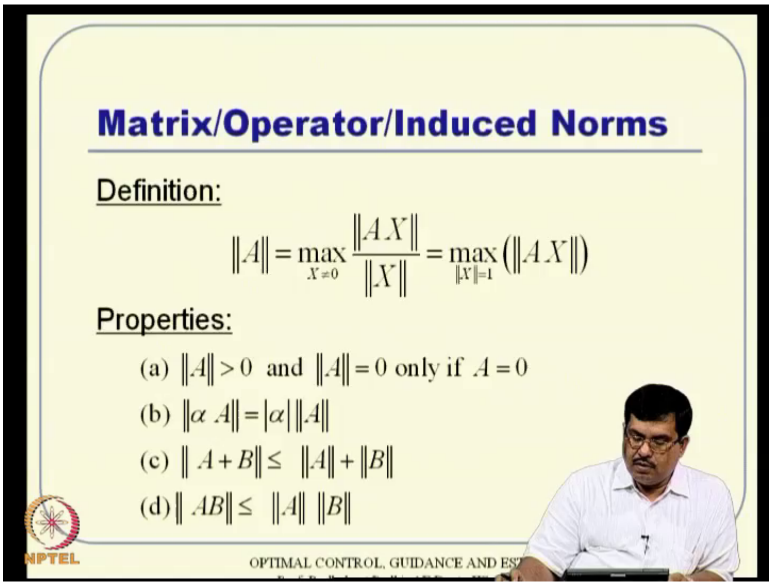
(Refer Slide Time: 33:35)



And some standard results, there are various norms which can be define computational purpose, and first thing that comes to your mind is one norm. So, that means you just simply take the components and then take the, kind of absolute values and sum it up, is nothing but one norm. Two norm is something very standard we all know that, and it is take of everything and then take off, and obviously here we can always argue that modulation not necessary true, but to have a systematic way of definition, it takes modeless everywhere, that is it actually. Then, if you have the norm three, you can do similar algebra; take third powers everywhere and then take one third, and similarly the pth norm in a standard notation since it can be defined that way. And infinity norm turns out to be like this, going by this definition and it can be soon it is an interesting a kind of result actually, that even though infinity to the power in one by infinity think like that are there, it is essentially, it has a meaning, and concept to be that the maximum value of this modeless of X I.

That means, if you simply take this modeless thing and do not edit up. Just look at the values and see which is the maximum value, and that value is nothing but infinity norm actually. And very interestingly it turns out that in finite dimensional vector space; that means the dimensional is not really infinity. In other words, X is some M components only; it does not contain keep on containing many terms, so infinity components really. Then in finite dimensional vector space, all norms of something like equivalent; that means, if you proof something using some norm, whatever norm is that, and some sort of a similar property will hold even if you talk about any other norm actually. So, this is standard norm property in finite dimensional vector space actually. It is not true in infinite dimensional space actually.

(Refer Slide Time: 35:29)



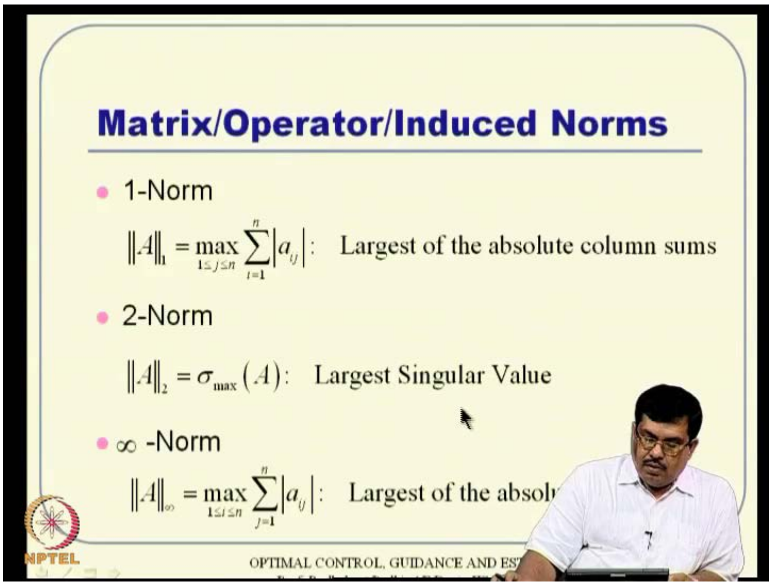
Now, going on, I mean moving on further, there is a concept called matrix norm as well, and the matrix norm turns out to be something like that. Again remember A times X is nothing but a vector. So, A times X is an operator basically; it takes transform vector X to vector Y basically, we have seen that little bit before. Now the question is if you, let us go back to that diagram, and here if you have A X and this is like a stretching operator. Even in general its vector transforms P 2 Q and there is a stretching operation involve there actually. So, in that context, the question here is, if I multiply any vector X with matrix A, then I am getting a vector y which is a different, I mean different magnitude and think like that. Now, the question is, how much does it (()), because of A, not because the, originally the vector itself is longer or shorter, we are not interested in that. So, originally it can be longer or shorter; then we are not interested in that kind of thing.

We are telling that, I mean, if I multiply a vector by matrix A, then I get some vector, and the question is how much I am stretching the vector, how much I am stretching vector, because I multiply it with matrix A, that is the question actually, and that can be argued, that I can evaluated this; that means I evaluate the norm of Y and divided it over again why norm of actually. So, this is something Y, something like that. So, that is how it is evaluated. Remember, this vector norm the both this numerator and denominator are vectors, and vector norms are already defined here. So, you have to pick up any norm; that norm, first norm, second norm third norm all that, and then you can talk about carrying out this algebra and it turns out to be the. It is actually define some sort of amplifying vector or amplifying property of this matrix actually. So, that is how it is defined as matrix norm actually. It is also called of operator norm, induced norm think like that, various terminologies actually.

It is also true that inside of carrying out this algebra if I constrained myself to this norm equal to 1, then obviously this is 1; norm of X equal 1, and hence, all that I have to see is what is the maximum value it takes actually. And remember, if there is a maximum operator involve actually; that means, I have to carry out something like I have to see familiar of X, I mean lots of X actually whatever is there, and then I have to do take a maximum value whatever happens to their. So, maximum applying property that comes out of A matrix actually. So, similarly instead of this algebra turns out to be little bit tougher, because I mean they will be infinite possibilities and all that. So, we confine ourselves to this constraint that now of X equal to 1, and then carry out something like a maximization operation, I mean maximization analysis and all that, and tell this is what value is actually. That is how it is define.

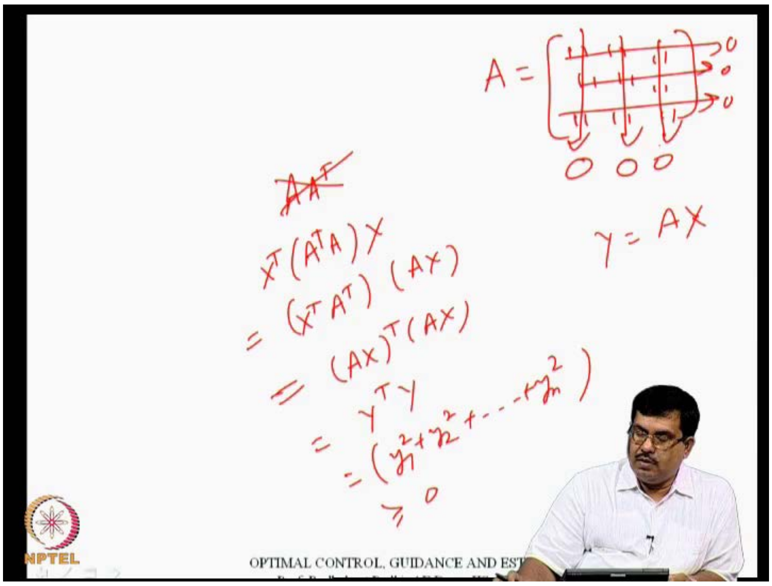
Again property since it will satisfy some of these properties so that norm of A is to greater than zero. It is zero only if the matrix A is nothing but zero matrix. All the elements of matrix zero, then it will happen to be like that. And for that properties B and C are extremely similar to what we saw in the, in the vector norm sense, and then it turns out to be additional property that, if I multiply A times B and taken norm, then turns out to be less than equal to norm A into norm B actually. This is a property that does not, it is not there in vector norm, but it is there in the matrix norm actually. So, I hope it is slightly clear now, that all that we are talking matrix norm is looking at the amplifying property of matrix itself basically, and it take how much maximum it can amplify for a vector X actually. So, that is the concept and these are some of the properties involved.

(Refer Slide Time: 39:22)



So, again computational sense, how do you compute. First norm; one norm if you tell, this is this definition; that means you take all the elements of the matrix, and then kind of take all the, I mean this absolute values, and then all that you have to do is, formulate this columns and all that.

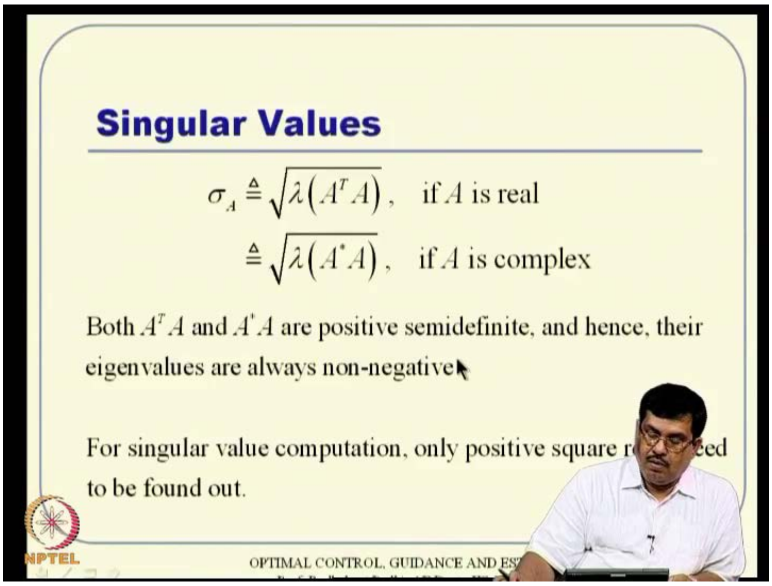
(Refer Slide Time: 39:48)



So, let me just, I mean just demonstrate that. So, this is a, this is something like if I have matrix A and then talk about something like entries everywhere. All that I take is just absolute values of everything, all the elements whatever I have, and then I can do a summation here, this way or a summation that way, either way. And then take, if I see the summations here and see whether out of these three, I mean three or whatever values are there, which is the maximum or I can also tell which is maximum here actually, out of all the three things. So, this is what the concept; one is the maximum rows sum and one is the maximum column sum actually. This is what is talked here, and it turns out to be the largest absolute column sums. It is your first norm, but takes column by column, and then see the values here, and then end of the columns actually addition of all the values, and then tell what is the largest value among those possibility; that is nothing but one norm.

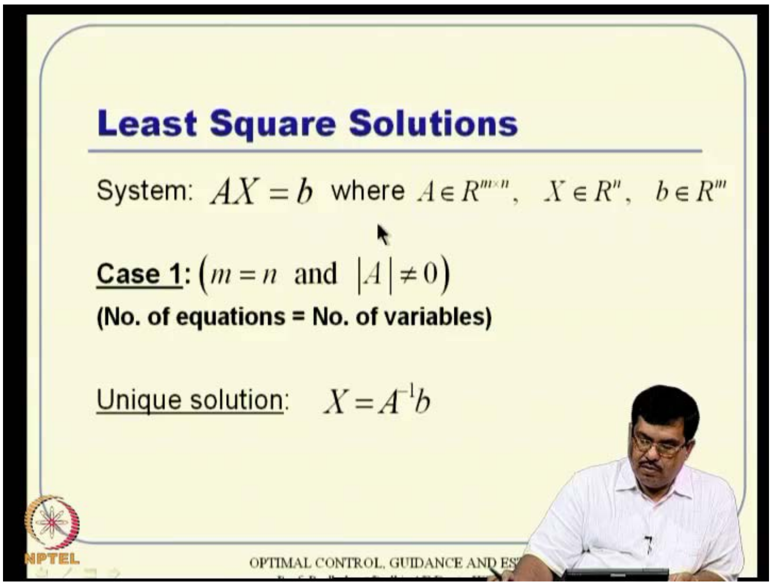
And similarly, infinity norm turns out to be the largest absolute rows sum as well; that means you do not carry out the column sums, but you carry out something like row sum, and then see what turns out to be in the right hand side, and then tell what is the maximum value out of that; that turns out to be infinite term, and just a hint of note, I mean how do you remember this two. Letter one turns out to be vertically, I mean placed. One turns out to be vertical notation, and then infinity turns out to be slightly horizontal notation. So using that, probably you can think up very (( )) algebra. But also we take a two norm many times in analysis and synthesis, and two norm turns out to be some concept called maximum singular value of it. Now, what is singular value, we will see that, and remember this matrix norm and all when you talk. The matrix need not be square; it can be, I mean it can be rectangular matrix as well actually, and singular values are also defined for rectangular values; eigenvalues are not by the way.

(Refer Slide Time: 41:44)



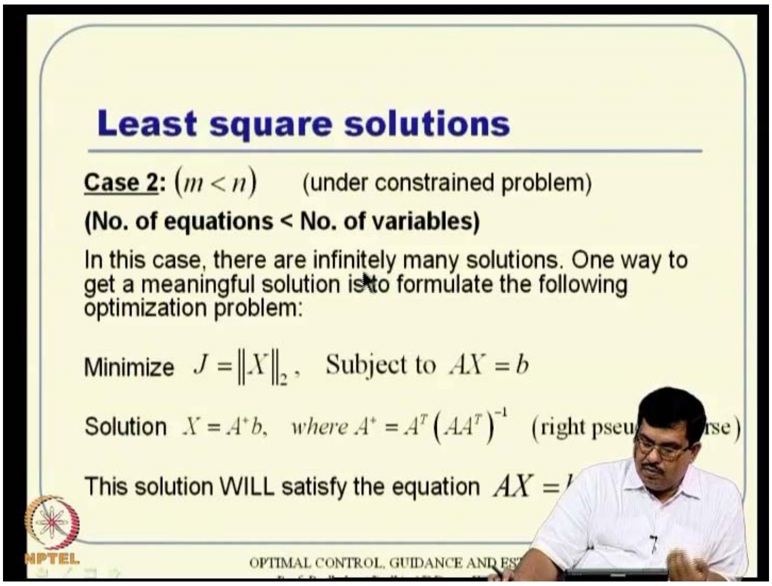
So, singular values can be defined like that. Remember A transpose is A always guarantee to be symmetric. So, you can talk about eigenvalues of those that matrix. Then remember A transpose A is also guarantee to be, I mean positive semi definite. So, the eigenvalues are guaranteed to be either zero or greater than zero. So, you can talk about a square root of that, and then take only positive square root of that and defined that is singular values. And in general, it is define some the like this. This A star when you see somewhere, this it turns out that, if your A matrix itself contains this complex numbers and all that, then this is the complex conjugate values and things like that, conjugate transpose actually. You take complex conjugates and then take a transpose. This will satisfy many properties similar to A transpose. If you simply take transpose, using this complex numbers and all, it will not satisfy. What we know in says in transpose real matrix actually. So, most of the algebra will be very comfortable, when you do this A star operation, it is nothing but you take the complex conjugates and then take the transpose actually, conjugate transpose. Anyway, so this is the idea of this, you carry out this algebra. Take the eigenvalues of this matrix; it is guaranteed to be positive and greater than equal zero. Take the square root of that when you land up with a singular values actually.

(Refer Slide Time: 42:58)



So, then there are concepts for least square solution and all that. So, the question here is something like this. If the, I mean there is an equation, linear system of equation AX equal to b, and then we are asking for solution for x actually. Now, in our previous knowledge and all that, it is very standard that if m equal to n; that means it is a square system; a square matrix, and we have equal number of variables is number of equations. Then there is a unique solution X equal to A inverse b. So, very standard result we all know that, but the question is, what is if A is non square matrix.

(Refer Slide Time: 43:39)

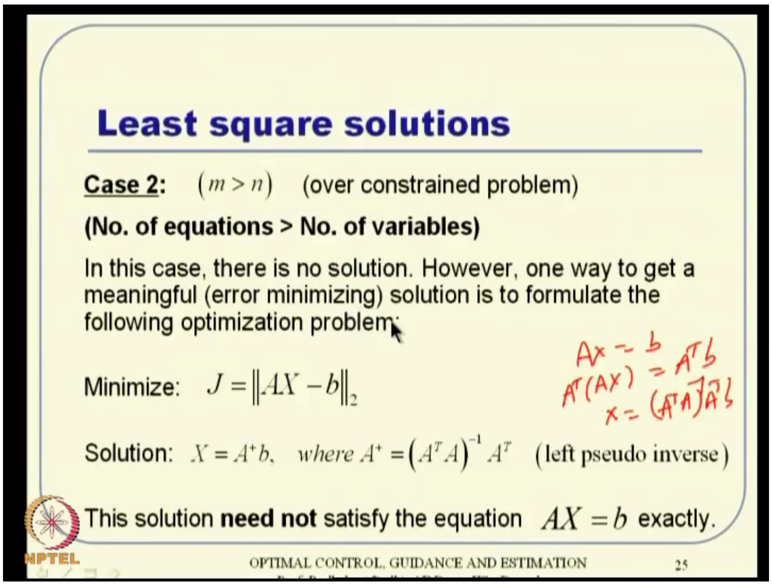


So, that is where, let see the standard results here, and if it is non square, then either m can be less then n or m can be greater than n. And here we are talking about m is less than n; that means number of equations are less than number of variables. Certainly something like an under constrained problem, and if its a under constraint problem, obviously we can think of having a solution, which will satisfy the equations exactly, and in addition to that we can do anything, something else actually, because we have more freedom and we have less constraints. So, we can think of a solution. In fact, it is infinite solution which will satisfy the equations exactly.

Now, out of the solution, that satisfies the equation exactly, what kind of solution that are of interest actually. So, obviously, one answer to that is minimum norm solution basically; that means, suppose you are talking about A X being a control variable for a second, then we want to have a solution for this, for which we have we will apply minimum control basically, because we are not compromising anything on the solution quality. Solution is certainly satisfied exactly, but we want to satisfy that with minimum control value as well actually. So, that is how you can think of the utility part of it.

So, that is how it is formulated, and the solution turns out to be something called pseudo inverse, x equal to A pseudo inverse b and then, it is something like pseudo inverse is defined like this, and also we know that pseudo inverse can be both right pseudo inverse or left pseudo inverse, and in this particular example, it turns out to be right pseudo inverse actually, so this is how it is. Then remember this solution will satisfy the equation actually as well.

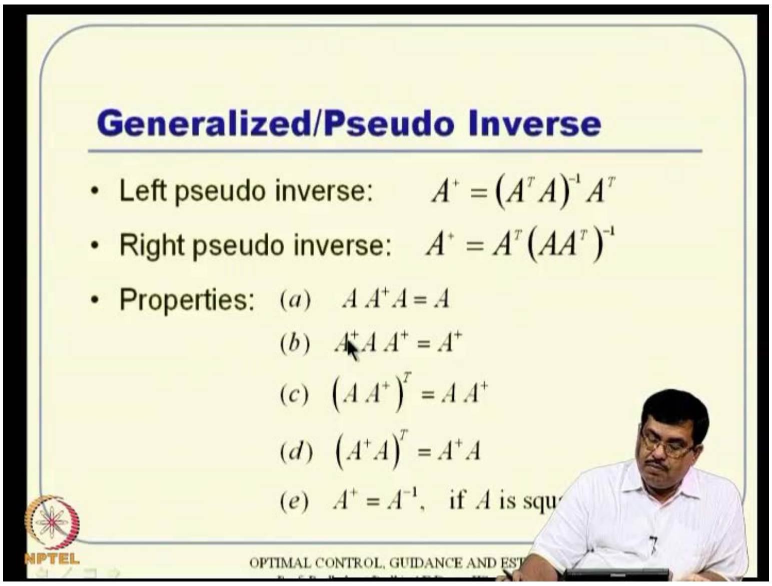
(Refer Slide Time: 45:20)



So, under constraint problems are typically luxury in a way basically. Then, what if this over constraint problem, exactly the reverse case, where number of equation are more than number of variables, and obviously in this case, assuming that all these constraints are linear independent; that means all this constraints have meaningful constraints, not some combination over the constraints and all that. Then it turns out that we cannot be a perfect zone. We can never aim to satisfy this solution exactly, but if you cannot do that, the next best is, how can you get a solution which will approximately satisfy this equation. So, that is the formulation here. It talks about minimizing the error quantity, second norm of the error quantity. Remember, A X equal to b is the equation. So, A X minus B is error sort of thing. So, if you talk about that, then it is nothing but minimization of the error associated with all the equations actually.

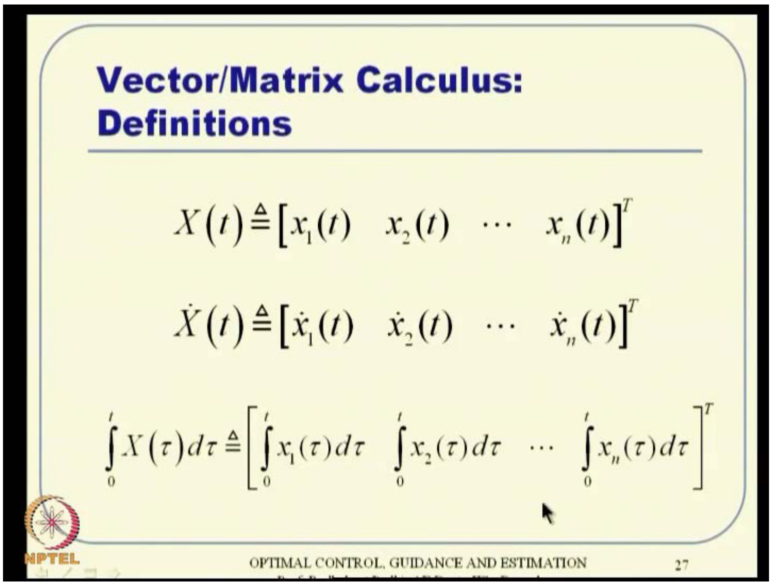
And interestingly, the solution again turns out to be pseudo inverse b, but this time this pseudo inverse is defining something like this. This is called f pseudo inverse, and sometimes it can be argued this way also, it is very easy to see that. So, if you have this something like A X equal to Bb then you all that you do is multiply a transpose both sides, and land up with that, and a transpose a comes from here. So, we land up with a transpose A inverse A transpose and B. So that is you can interpret that way as well. So, easy to see that actually, but you will not be able to do that exercise that way that easily in this case actually; it requires little bit further algebra actually, but the point to note here is, this solution what you are getting here will not satisfy the equation exactly. It will satisfy only approximately.

(Refer Slide Time: 46:54)



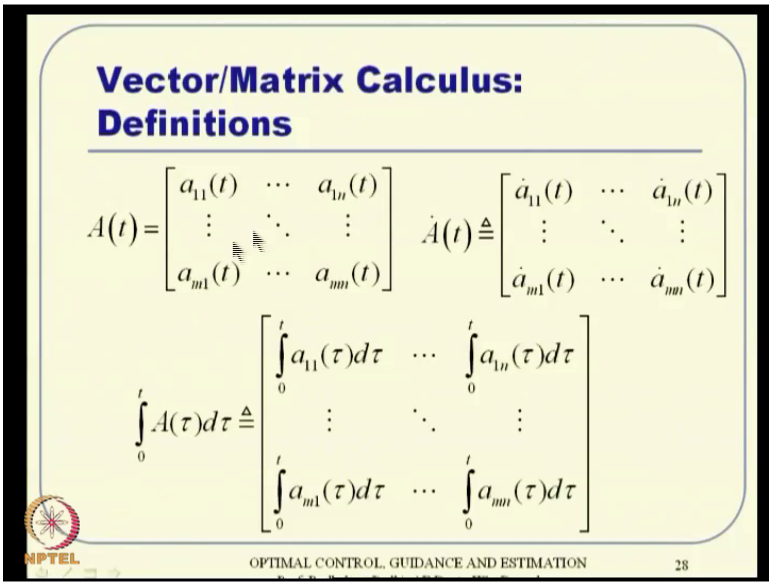
So, this in general, this is the left pseudo inverse; this is the right pseudo inverse and some of the properties that pseudo inverse satisfy something like that, and you can see very closely it will feel as if it is A inverse actually. If I talk about A pseudo inverse A, that means if I close the bracket here or close the bracket here, it feels like identity actually. So, that means it is a. Similarly, if I take the other operation, if I say this bracket or close this bracket, and interpret the pseudo inverse as a nothing but inverse, then I will land of with other one actually. So, it will feel very close to what the inverse properties satisfies. That is why it is called pseudo inverse actually, and also remember in a compatibility sense, if A happens to a square matrix and determinant of A is not equal to zero, then pseudo inverse is nothing, but inverse actually, and it is very easy to see that here as well. So, we will land up with matrix inverse sort of thing actually.

(Refer Slide Time: 47:53)



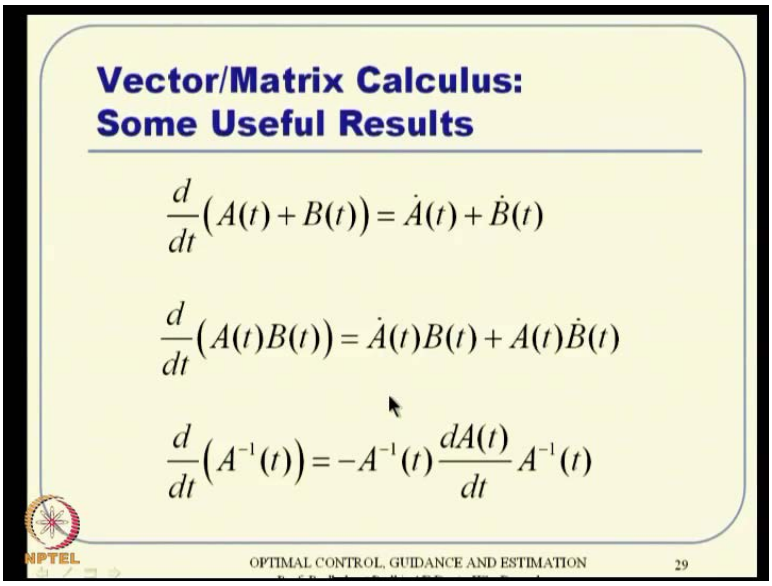
Now, before we conclude this lecture, it is also time to see some of this vector matrix calculus in general. So far we have been talking about vectors and matrices containing numbers, but what if they are functions of some other variable. So, that is where our most interest lay, I mean lies as well, but most of our state variables and control variables all these are time baring actually. So, if you have a time baring function setting as a matrix, and then you can certainly talk about their differentiation, as well as their integration, and by definition, it is simply turns out that if I take component by component differentiation, that is my differentiation of vector X, and if I take component by component an integration, that turns out to be my integration of vector X actually.

(Refer Slide Time: 48:38)



So, that is simply by definition, but there are interesting algebra after that. So, this is also true for matrix actually. If I have matrix A with all the elements are time baring, and by definition, D by DT of A is nothing but all these, the component by component differentiation. Similarly, integration is also component by component, very easy to see that, but after that there are certain standard results again. If I talk about differentiation of A plus B; where Aa is a function of time and B is a function of time. Then it is nothing but A dot plus B dot, that is how it turns out to be. It is very easy to see that anyway. So, all that you have to do here, is again go back to this definition; A is nothing but like this, so A dot is like this. Similarly formulate B is nothing but B 1 1 B 1 2 like that. So, B dot will like that.

(Refer Slide Time: 49:31)

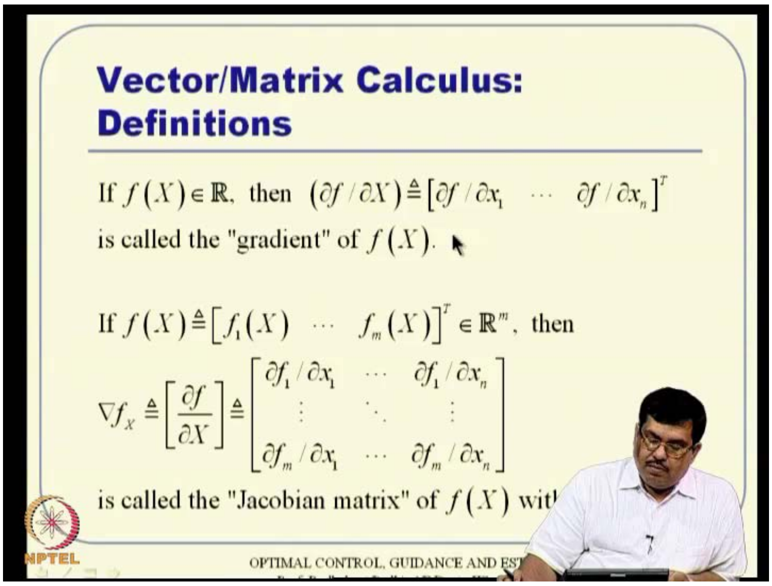


So, add it off. One side you add of A t plus B t, and then, take differentiation, use that standard definition. Other side, you just add it up this A dot and B dot on the right hand side. You will very easily see that this is very much true, and it is feels very much trivial as well actually in a way. So, d by dt of A plus A t plus B t by definition. Simply you expand the entries and then put the definition or differentiation, combine it back again and you will very quickly see, that it's nothing but A dot plus B dot. Similarly, but this one is very interesting, and remember, d by dt or the differentiation of A times B; where A and B both are functions of time.

It is again satisfies the result that we start very closely know in scalar algebra actually, scalar calculation, but the only difference that careful thing that you have do here, is you cannot change the sequence really. Remember, AB is equal to is not equal to BA in general. The dimensions can be different, way can be define and think like that, and A dot and B dot will carry the dimensionalities a and b along with them; simply form definition dimensionality does not sense. So obviously, these definitions are reverse are and all that is define actually. So, be careful while using this formulation, I mean formulation all that. It is true, but it is true in the since of sequences only. And this is also easy to kind of derive that d by dt of a inverse is nothing but that.

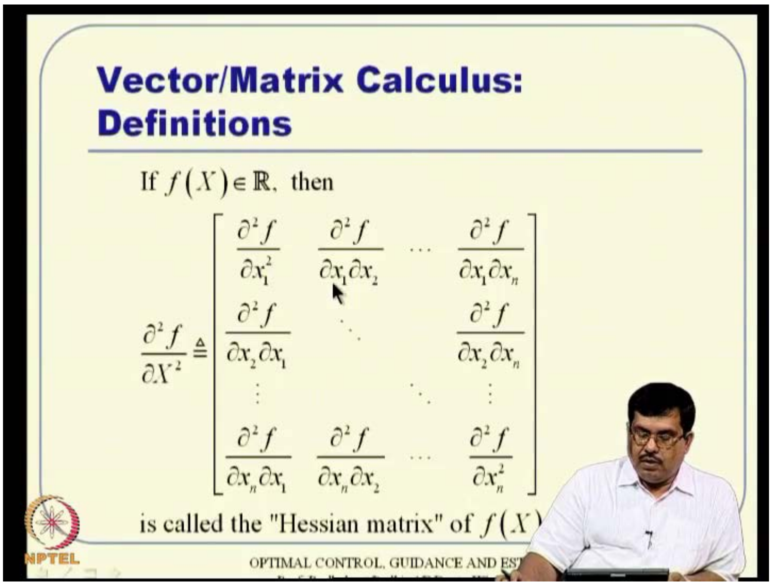
It is again if you just interpret a scalar quantity, then it is nothing, but A dot divided by A square actually, but in a matrix sense we have to write that only. We cannot write that too, and also remember division of matrices are not defined. So, da by dt is a matrix divided by somebody writes a square all that. So, this matrix divided by another matrix is not define actually; it is simply the multication that is defined actually. So, it turns out to be like that, then using this theorem, I mean using this result, it is easy to solve that. Just that you have do is, take B equal to A inverse and then carry out this algebra, it will land up here actually. If it is B is A inverse, then this is identity, so d by dt is zero and then right hand side algebra we carry out, then it will very easy to solve, so that it turned out to be like that.

(Refer Slide Time: 51:36)



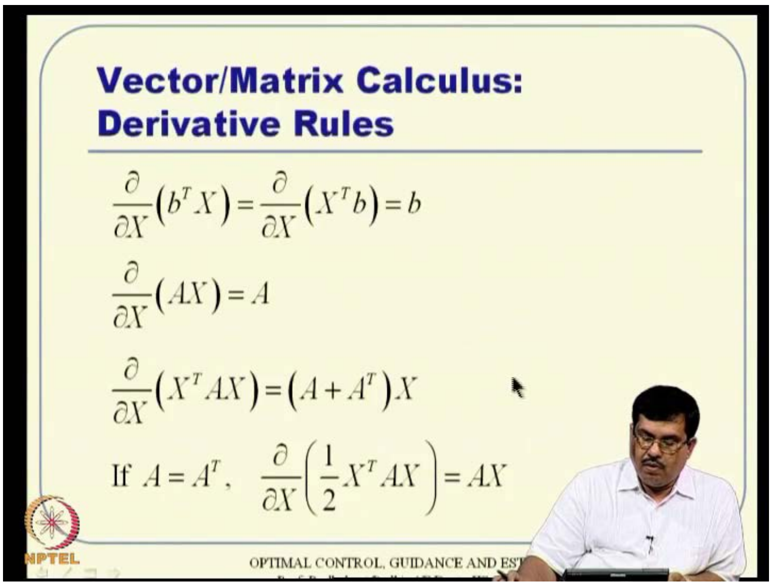
Anyway, so now moving further and calculus, vector matrix calculus. The gradient vector is something that heavy heavily use in optimal control theory, and the gradient is define something like that. Again depends, see the x can be n dimension, but function of x let us assume that it is scalar quantity. So, it is very standard example its something like x 1 square plus x 2 square plus x 3 square like that actually. So, ultimately the quantity is scalar quantity, but x vector contains components basically. So, in that sense, the del f by del x is a gradient vector, and standard x commonly known as and that is defined like that, remember, this is a transpose. So, that is actually by definition, we have seen that actually kind of column vector actually. This is gradient vector denoted by del f by del x. It is a column vector with partially derivatives like that. However, if f itself is a vector; that means X is a vector and f is also a vector, then the gradient vector is not called really a gradient; it is called Jocobian matrix, and this is define something like that. We just see that in something like linearization of non-linear system and all that. See you have to carry out all these partial derivatives and then put it in this form, then you will end up with Jacobian matrix f of X.

(Refer Slide Time: 52:57)



Then it is ok. However, second derivative is, if f of x is a scalar quantity, then second derivative is define like this, and remember, this is symmetric matrix, guaranteed to be, is called in hessian matrix actually, because this quantity del square f by del x 1 and del x 2 is nothing but this quantity actually. So, no matter what order you take both are same.

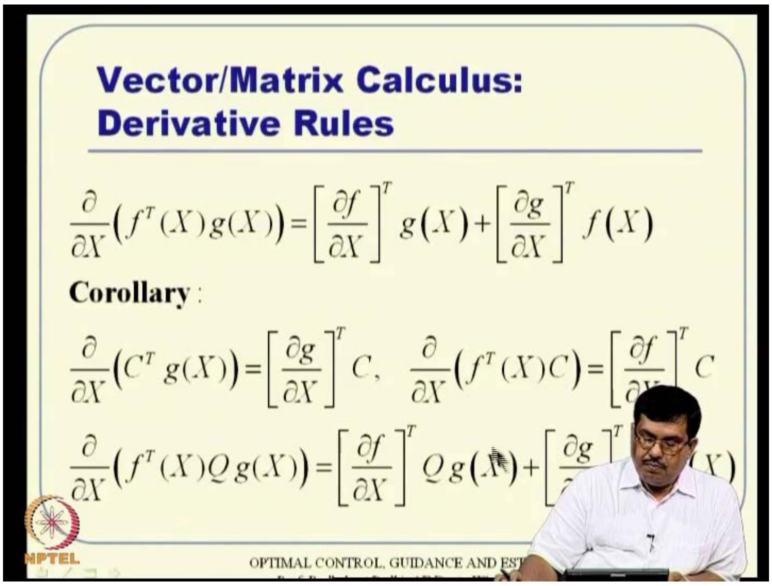
(Refer Slide Time: 53:23)



So, using that property, it turns out to be this matrix is kind of symmetric matrix all the time. Then, there are certain derivative rules and all that. See, del X, if B is constant vector rather and x is a variable, then del by del x b transpose or del by del x transpose b, it turns out to be b actually. You can think of it is a kind of a scalar algebra sort of thing, but in general, in calculus, I mean vector matrix calculus also, this property is true. It is very easily to, so that again simply by carrying of the long and algebra. You take v equal to v 1, v 2, v 3 upto v n.

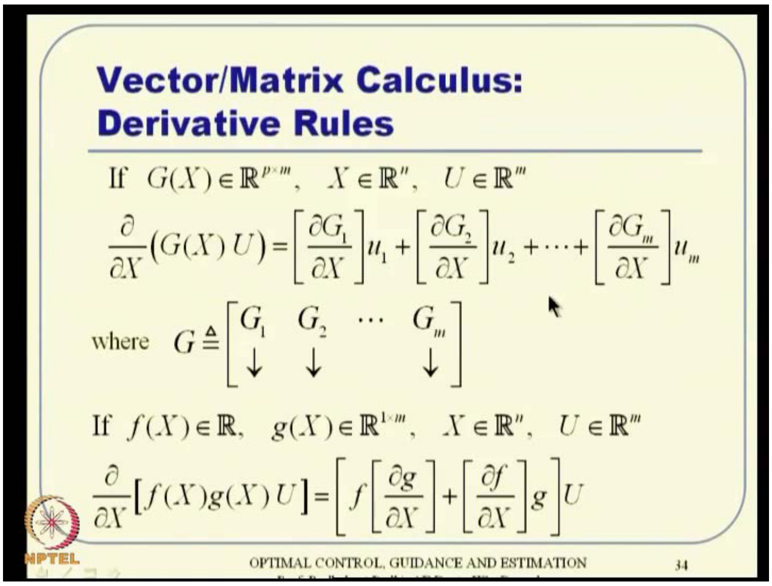
Next, x 1 x 2 x n. Formulate these and take the derivative component, use the definition of derivative and then land up that it is nothing but, V1 V2 V3 arranged in a column actually. So, it is nothing but B vector. Similarly, interestingly turns out that partial derivative of AX with respect to X with turns out to be A, and partial derivative of X transpose a transpose to be like that, and as a corollary if A equal to A transpose; that means this is there. Then I can always talk about del by del X del X of half of del X nothing but AX. Just use A plus A transpose is nothing but 2 a and it will get cancel out, and this half will cancel out with a X, and this relate to something that will use it later. In general, this is true, but if A is symmetric, then this is true.

(Refer Slide Time: 54:45)



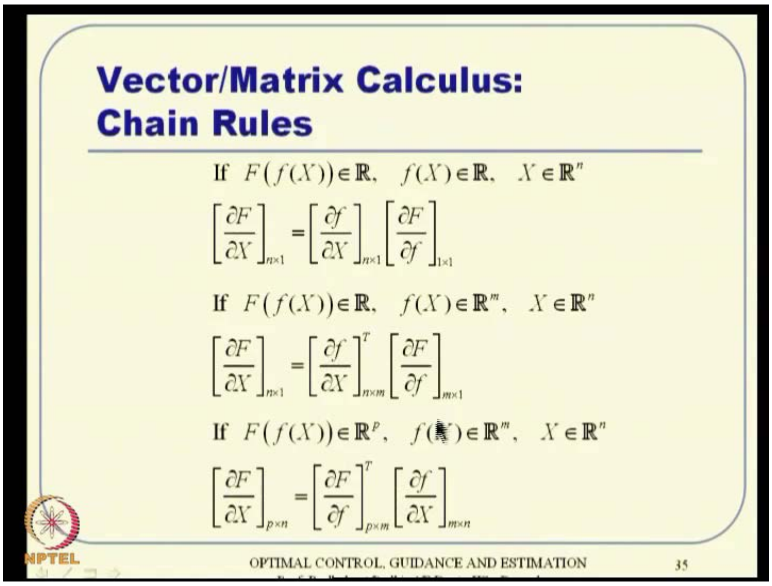
Then, there are further results, which tell you what if there are two functions like that. So, del by del X f of X transpose g is what, and again satisfies very close relationship that we know in scalar calculus, but again you cannot change sequence here. So, this result is given something like this; del f by del X transpose into g of X plus del g by del X transpose f of X. And as a corollary if f happens to be C vector, then it turns out to be like that and think like that actually. These are standards coming out of that actually. So, remember f and j can be vectors by themselves actually here. So, many times you use this result also. Especially when we talk about this control f fine system and optimal control with that and think like that the way.

(Refer Slide Time: 55:35)



Now what if g f X is a truly matrix, then what actually, and then turns out that if I take a g f X times U, this is where this control over f fine system and all will come. I will tell about that as you go along the course, what is f fine systems and all that, but there will be a terms something like that X dot f of X plus g of X times U, that is, that kind of form actually, phi f g X times U and then you have to carry out these algebra. That is necessary for doing this optimal control analysis and all that, so then this result of turns out to be like that. And this can be derived like that where g 1 and g 2 and all vectors, are column vectors coming out of this g matrix. Then you take the partial derivative and put it there in this form actually. Again these are longer algebra; it takes about a couple pages, but you can just would the plugging the standard results from definition and carry out longer then algebra, you will get it there actually. So, this is special case, these things do happen, I mean if you take only, this is scalar, this is vector something like that, and then this turns out to be like this. It is corollary of all these, that we know.

(Refer Slide Time: 56:35)



So, what about same rules then, I mean if you have function of function, then what actually. So, it terms out to the X is vector, but f is function and then it is another function of that, and then the game starts that X can be a vector in general, but this small f and big F, You can take different forms. First, you take if the small is f vector and big F is also vector, small f is scalar and big f is also a scalar, then this is the result. And if small f is scalar but big f is a, I mean small f is a vector but big f is a scalar, then that is the result, and if both are vector, then that is the result. It is very similar, I mean similar looking, but worrying about which comes first, and whether there is a transpose or not actually. And again, these are very standard reason. You can have it in somewhere, and then, if you want use it, you can use actually. So, that is how I can summarize this, this, lecture that that, we have got some over view of what is non-linear system, and then we carry out what is, I mean what is called linearization systems and all that. Then which I talked about various matrix properties that goes along with those matrix actually, and many of this things, we will use it subsequently in our further discussion of topics related to optimal control. With that, I will stop here. Thank you.