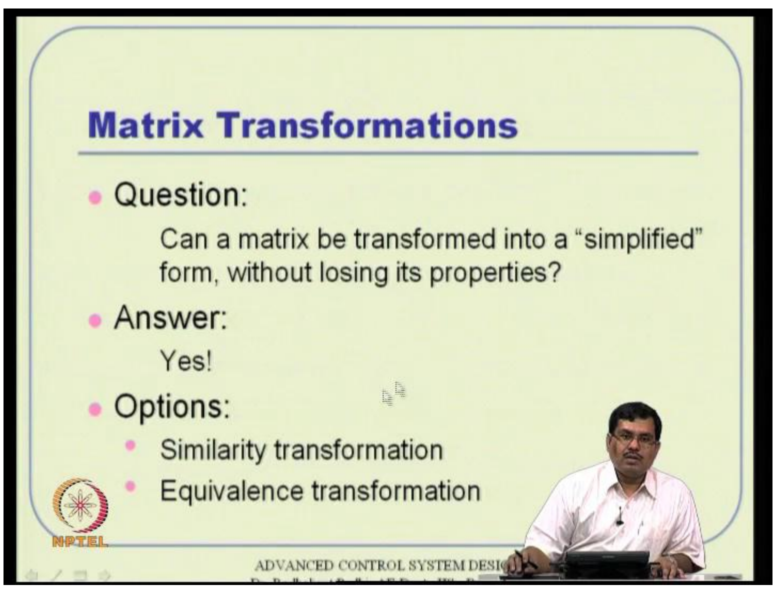
Advanced Control System Design
Prof. Radhakant Padhi
Department of Aerospace Engineering
Indian Institute of Science, Bangalore

Lecture No. # 13
Review of Matrix Theory – II

So, we have the seen many of the properties something like basic from starting from basic definitions to Eigen values and Eigen vectors, next and then Vector Matrix norm to like concepts something like induce norm and then we followed up with these linear equations solutions and thing like that, and ended up with Pseudo Inverse Concept and things like that. Now, we will continue further of some of the topics that are useful in modern control theory in this particular lecture.

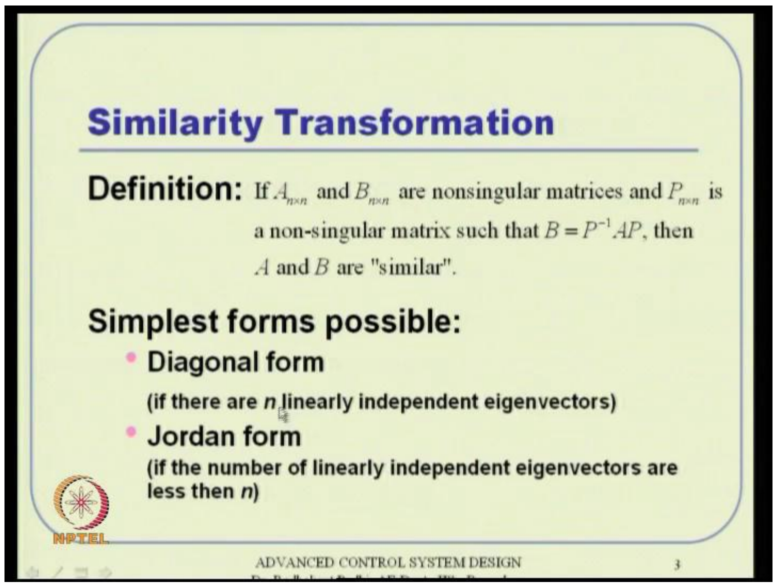
(Refer Slide Time: 00:43)



So, the first thing that comes to mind is something called Matrix Transformation. So, this Matrix Transformation is the question that we are asking here is something like this. Can we can a matrix, suppose we have a matrix, can this matrix be transformed into a simplified form without losing its properties actually.

I mean the properties of the Matrix the basic properties of the Matrix we do not want to lose, but still suppose a given Matrix say I want to transform it to a Matrix B such that if necessary I will recover back my Matrix A actually from B actually and this answer to this question is fortunately, Yes. And the options that we are having is like Similarity Transformation and Equivalent Transformation these are all things that we study in this particular class actually.

(Refer Slide Time: 01:40)

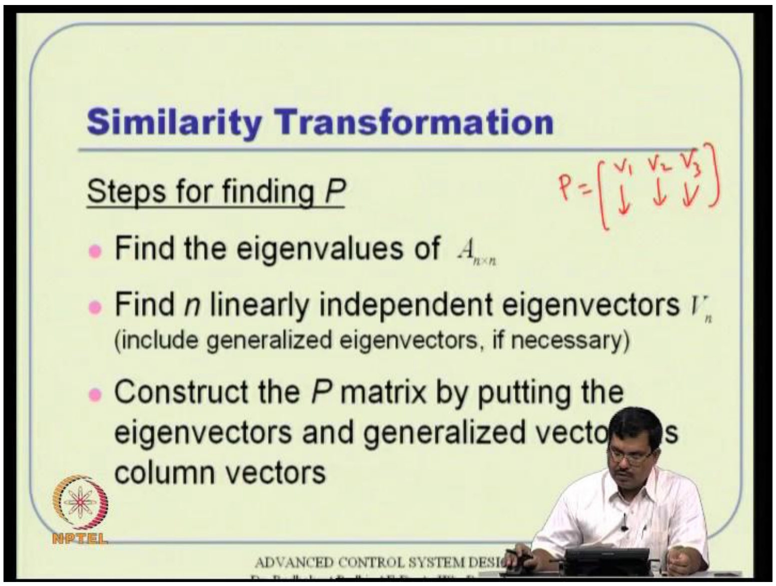


So, what is a Similarity Transformation first of all? So, if a I mean Similarity Transformation is a concept for square matrices and Equivalence Transformation is a concept for non square matrices as well. Now definition wise, if A and B are n by n non-singular matrices and P n by n is also non-singular matrix, such that this relationship holds good; that means, B is equal to P inverse AP, then we call A and B are similar actually that is the definition.

So, A and B has to be related to each other with this P inverse AP, and all that is necessary is that P matrix has to be non-singular that is all (( )). So, I can really talk about P Inverse whenever I need actually. So, if you do this transformation; let us say we somehow find this P matrix actually, and then carry out this P Inverse AP, then it turns out that this matrix A can be reduce to like a diagonal form or in if diagonal form are not possible, then it is a Jordan form will end up with actually.

And then, this Diagonal form is possible provided this Matrix A as n linearly independent Eigen vectors. Remember, A is n by n Matrix. So, it can have in Eigen Vector n Eigen vector and if this a Matrix has n linearly independent Eigen vectors, then it will be lending up with A B Matrix which is purely diagonal. Otherwise, if you does not have n linearly independent Eigen vectors, which happens provided this some of this Eigen values of the Matrix are repeated, there can be a chance of getting general Eigen vectors and thing like that, then we will end up with a Jordan form which is very close to a diagonal form, but it is really not a diagonal form, we will see that actually.

(Refer Slide Time: 03:25)



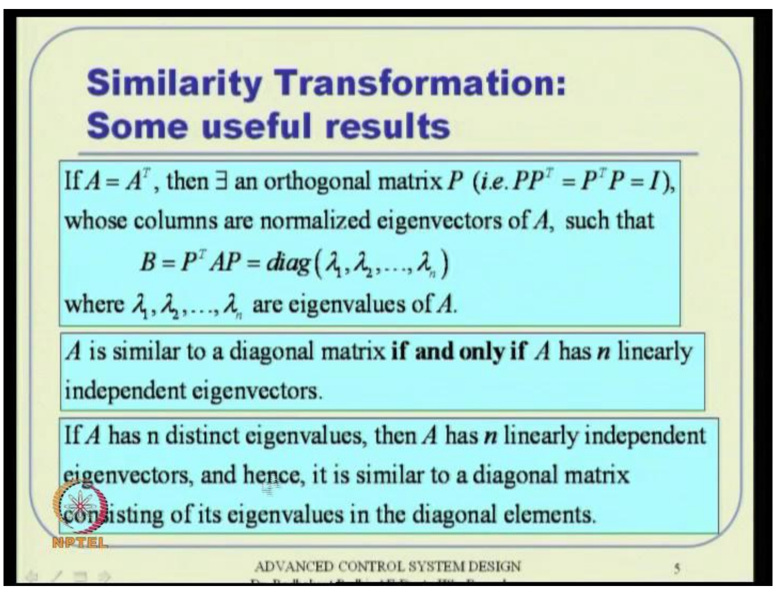
And steps for finding this P is how do you find out this particular P? These steps are something like this and find this Eigen . First of all you have to find the Eigen value Eigen value Eigen vector pair for this A Matrix actually. So, you find the Eigen values first and then go ahead and find the n linearly independent Eigen vectors first, and then if the Eigen vectors are not possible you can include generalized Eigen vectors. Remember, these are like Eigen vectors so that means, generalize Eigen vectors are guaranteed to be linearly independent compare to the regular Eigen vectors actually. So, you go ahead and find n linearly independent Eigen vectors, then you construct AP Matrix by putting this Eigen vectors side by side actually in terms of column vectors. So that means, V 1, V 2, V 3 then P

Matrix is nothing but V 1; simply I put it the first column; V 2 I put is a second column; V 3 I put it a third column that is V Matrix actually.

And really does not matter what magnitudes you take, because the operation that you see is P inverse AP. So, this may this inverse and P inverse will and multiply by P will take care of that magnitude actually. They are not so much concerned about normalizing that that Eigen vectors actually. If you want you can normalize, but there is no particular necessity of doing that actually. You can simply put the Eigen vectors side by side and if suppose V 2 and V 3 are not really independent vectors, then V 3 can be a generalize Eigen vector that we discussed last class actually.

So, that is how we will get a get the P Matrix, then you construct this B matrix that way P inverse AP that is all you do actually. So, that is what will give us like P matrix that you want actually. And we will not take very specific thing because these things this is just a review class anyway, so it not a matrix theory class for say. So somehow these things I will assume that you can find suitable example from text books actually.

(Refer Slide Time: 05:21)



And let us say some other properties very useful properties rather. Then, if a matrix is symmetric; that means, A equal to A transpose, then there exist an orthogonal matrix P.

Remember, orthogonal matrix means inverse is transpose; that means, PP transpose P transpose P is equal to identity. So, what it tells? If a matrix is symmetric, then there exist an orthogonal matrix P which columns are normalize Eigen vectors of A right such that B equal to P transpose AP. You do not have to talk about inverse anymore remember that, because P inverse is P transpose anymore; I mean P inverse is P transpose in this case.

So, B is nothing but simply P transpose AP so that the algebra is much much simplified actually. So that happens if you have a symmetric matrix actually. Remember just because just by having a symmetric matrix you have such a beautiful property actually. If you have symmetric matrix, then it is guaranteed to have an orthogonal Eigen vectors and then the P matrix will turn out to be orthonormal; I mean orthogonal, because if you put that Eigen vectors in normalize that. So remember a orthogonal matrix the columns of the orthogonal matrix needs to be orthonormal.

So, in this case you need to formulate A matrix P with the components of P that columns of P should be orthonormal to each other; that means, Eigen vectors has to be normalize first actually (( )). Once you normalize the Eigen vectors put them side by side, formulate the P matrix then P transpose is P that is all you do actually you do not need for the algebra actually.

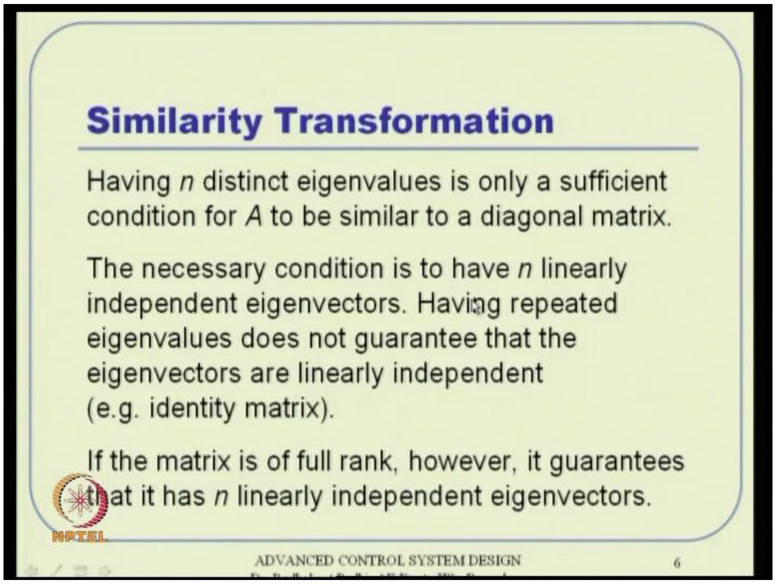
The next next property is a is similar to a diagonal matrix. If and only if A has linearly independent Eigen vectors. Remember that is the thing you are talking about the; I mean, we are always watching out looking out for whether you can really reduce it to a diagonal form, that is a simplest form that you can have actually. So, that is possible if and only if this A matrix has linearly n linearly independent Eigen vectors actually. It will not dictated by Eigen value condition, it is dictated by Eigen vector condition actually.

If a matrix has n linearly independent Eigen vectors, then it is possible to happen. Remember, if the Eigen values are non repeated it is guaranteed to happen. If it is repeated, then there is a chance that it can still happen, and then there is a chance that it may not happen I mean, so that there is the real requirement is that you should have n linearly independent Eigen vector, that is the real requirement actually.

The third property; if A has n distinct Eigen values obviously that is what I just told, the the the non repeated distinct Eigen values, then A matrix is guaranteed to have n linearly independent Eigen vector and hence by this property too, it is similar to a diagonal matrix and on the diagonal it will contain Eigen values actually. The diagonal matrix that the B matrix is guaranteed to be diagonal, but the diagonal matrix is nothing the Eigen values actually on the diagonal. What you have in the diagonal elements same Eigen values will pop up actually.

So even without carrying out this algebra, you can simply take what is B actually. If you simply find out the Eigen values and the similarity transformation, you do not have to do all this algebra, you simply tell that is may be matrix actually.

(Refer Slide Time: 09:00)

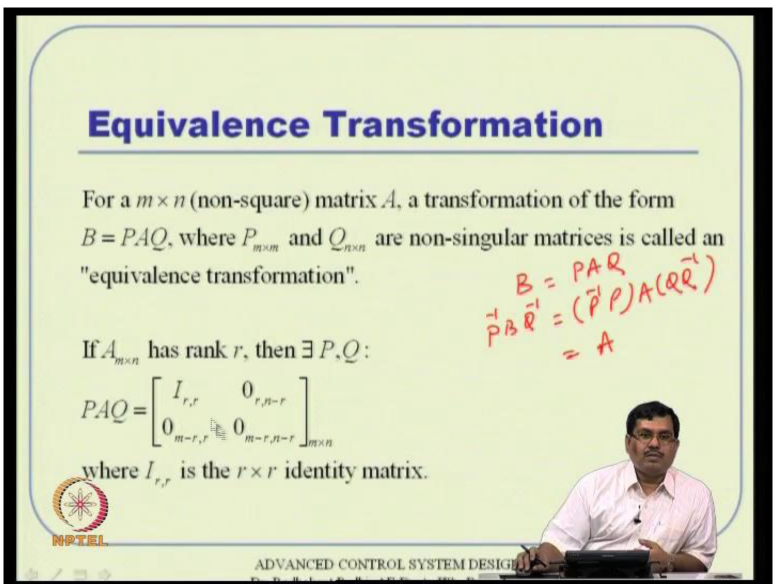


So, similarity transformation we will we will continue for the discussion what it tells us? It tells us that having n distinct Eigen values is only a sufficient condition. The necessary condition is actually n linearly independent Eigen vectors.

If you have linearly independent Eigen vector n of them that is actually the necessary condition and this actually comes from the sufficient condition, because if it has n distinct Eigen values, then other one is guaranteed actually and hence it is the result is very obvious actually. 0 0 0 and then, 0 1 0 0 0 and then, 0 0 1 0 0 all that all those vectors are actually linearly independent of each other. They are all orthonormal orthogonal to each other as well as normal. So, there entire identity matrix is actually a orthogonal matrix actually.

So, what does it tell you? Having repeated Eigen values does not necessarily guaranty that Eigen vectors are linearly independent actually. Can be linearly independent, it cannot be need not be actually. And now, the next question is if the Matrix is a full rank? That means, the rank of the matrix is n, it guarantees that the n linearly Eigen vectors will happen actually. If you have a full rank matrix that n by n matrix has a rank n, that is actually we guaranty that will have linear n linearly independent Eigen vectors and hence you can actually reduce that matrix to a diagonal form actually.

(Refer Slide Time: 10:34)



Now, the next concept is equivalence transformation. We will go ahead and see what is Equivalence Transformation now. See all this similarity transformation is very nice and further examples and all you can see it from the books; I mean this is very apparent from small discussions like this, but if further examples you can see from books actually. Now the question is, if you have an non-square matrix rather m by n matrix, then obviously you cannot talk about Eigen values, Eigen vectors. These are all Eigen value vector pair is good only for square matrices actually.

So, if you have a non-square matrix how do you reduce that? Now, it we actually now we will ask the different question with we will not constraint ourselves to like P inverse AP sort of thing. So, this we will we will just tell this can we multiply pre-multiply by P and post-multiply by Q another matrix actually. Where the condition is P and Q are suppose to be non-singular matrix and that is all. P and Q are suppose to be non-singular that is all actually.

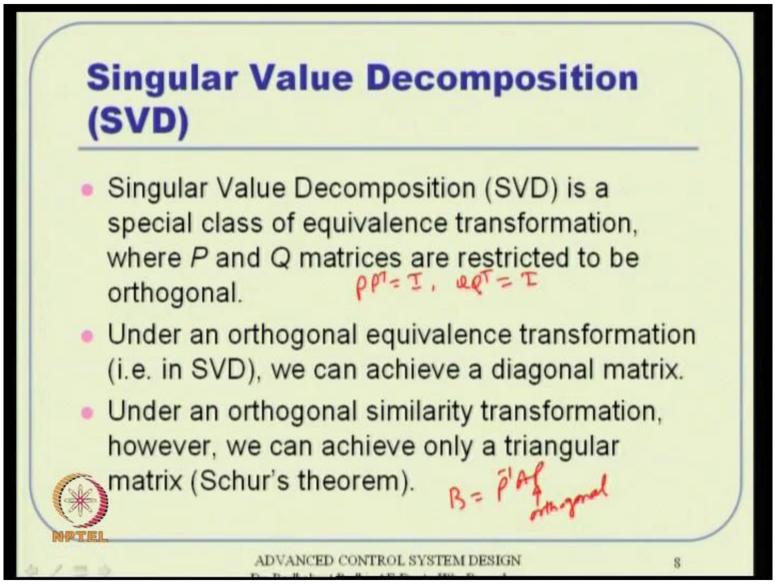
So, why that is necessary? Also remember, that I want to recover my A matrix; I mean from the B matrix also. If I have B equal to PAQ I will I will recover it by multiplying this P inverse B Q inverse also. This will be P inverse P A and QQ inverse actually. So that is nothing but A matrix; this is I times a times I that is a matrix. So the real requirement here is that P and Q has to be non-singular; otherwise, I cannot talk about their inverse here actually.

So, if I take two different matrix P and Q they are suppose to be non-singular. And then I carry out this algebra P times A Q then I will end up with B matrix, that is called equivalence transformation. Now, it is so happens that if you really do this algebra and just leave (( )) actually; that means, P and Q are just non singular nothing else actually. Then, it is so happens that you reduce it too much actually; that means, the reduce form can always take you to this kind of a form actually. You will simply have let us say the matrix have rank r, then you will left out with an I matrix identity matrix r by r and all other elements of that matrix B is going to be 0.

That does not contain too much of an information actually. So, I do not want to lose too much of information by doing so much of by bringing in so much of flexibility actually. And if that is the case, I want some diagonal elements to pop up, some elements to be there instead of identity here. And I can do that by constraining this P and Q little more actually. The only constraint right now is that P and Q are non-singular, but let me constraint a little more and ask, can I now talk about something like a orthogonal P and Q basically.

So, if I talk about an orthogonal P and Q then that is going to happen; that means, let us see little down the line if I orthogonal P and Q means, what it is P inverse is P transpose and Q inverse is Q transpose? So, instead of talking about just a non-singular matrix I will talk about orthogonal P and Q also, then it will happen that these diagonal elements are no more 1. Some numbers will pop up and those numbers are nothing but what are called as singular values actually. We will see that slightly down the line.

(Refer Slide Time: 14:13)



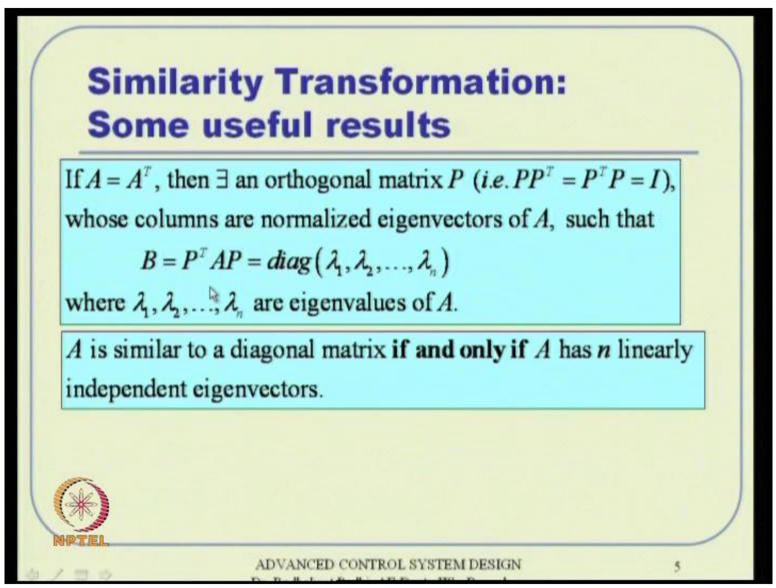
So, that takes us to Singular Value Decomposition and this is Singular Value Decomposition SVD popularly come a short form and in fact matlab has SVD command also. If you give a SVD within brackets some matrix that will give you a Singular Value Decomposition; it will give you both P Q and then B matrix everything it will give you actually. That is the command in matlab also with you. What is it actually? The Singular Value Decomposition is a special class of equivalence transformation, where P and Q matrices are restricted to be orthogonal; that means, PP transpose equal to QQ transpose equal to identity what you can do that really actually. So, you have to write PP transpose equal to I and QQ transpose equal to I. I cannot make them equal to that this is no more equal to each other, because the dimension of P and Q are not same. So that the I matrix that you are talking here and the I matrix you are talking here are of different dimension. So individually they are, but I cannot make it equal to here actually just keep that in mind.

Anyway so, if I constrain these matrices matrixes P and Q, PP transpose equal to I and Q Q transpose equal to I, then the the the reduce form that I get actually is nothing but Singular

Value Decomposition. And Then, I will we will see that actually in an example also. So, what it tells actually? Under an orthogonal equivalence transformation orthogonal equivalence transformation, we can actually achieve a diagonal matrix provided we start with a diagonal matrix actually. So, this this particular thing is actually in general it can be a non-square matrix. This is the reduced form is m by n actually. The dimension remain same for n V anyway.

So, if I if I really have to do that, then if I start with a diagonal; I mean, if I start with a square matrix then actually we can go and reduce it to a diagonal form also by Singular Value Decomposition. Really do not have to do this equivalence transformation. You can also do singular value decomposition for square matrix as well, this more general concept actually. So under an orthogonal Similarity Transformation however, we can achieve only a triangular matrix which is that is called as Schurz's theorem actually.

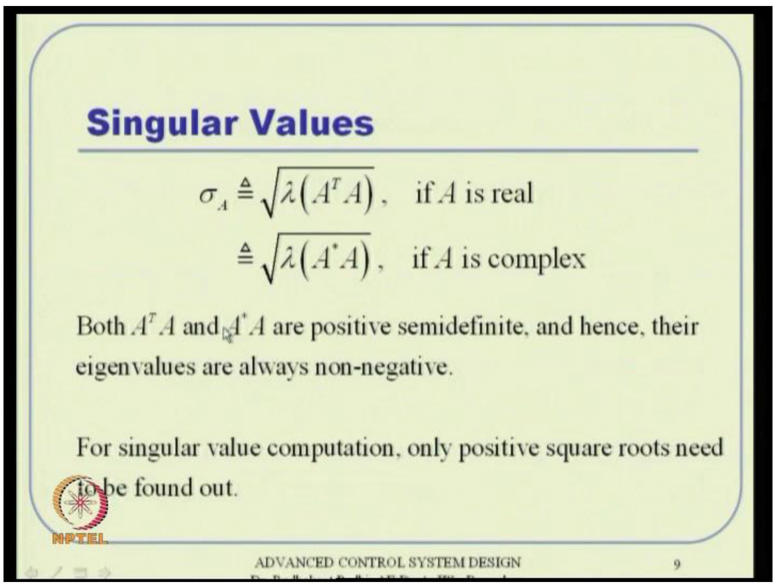
(Refer Slide Time: 16:45)



Remember, similarity transformation does not constraint this as orthogonal matrix right. It happens only when there is a there is a; I mean, this is symmetric matrix basically. In general it is not symmetric, we do not talk about an orthogonal matrix here actually right. You just simply talk about a matrix which is something like P inverse AP that is all we talk about that not more than that. But here, you talk about more than that and we will talk we are actually constraining P and Q to be orthogonal matrix by purpose actually. We are just doing that purposefully.

But if you do that also in the similarity transformation; that means, in that whatever you are talking this B equal to P inverse AP in that it constraint P to be orthogonal actually; this P has to be orthogonal, then what happens is in general you may not be able to find a diagonal matrix actually. That happens you will still be able to find a diagonal matrix provided A is symmetric in that case probably yes, but in general it is not. Then in general, you will find if this B matrix will happen to be a triangular matrix actually. So in general, you can go to a triangular matrix not more than that actually just keep that in mind actually.

(Refer Slide Time: 18:04)

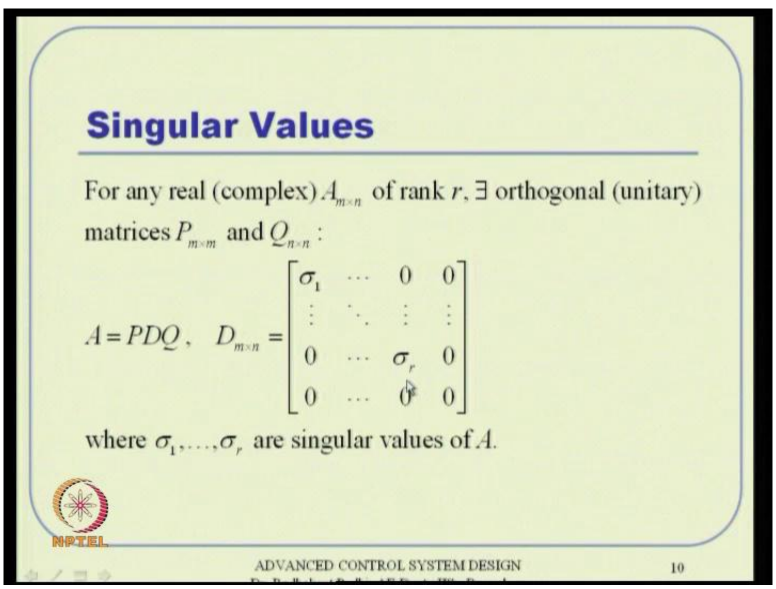


Now, let us say while let us define singular values formally and singular values can be defined something like this; positive square root of Eigen values of A transpose A if A is real and there is a equivalence if A is complex. Most of the time in control theory we will deal with A is real matrix anyway. But in general, if A is complex matrix, then you can also define something called A star which is like complex conjugate transpose. This is simply transpose, but this is actually complex conjugate transpose.

So, concepts are very parallel; suppose whatever you do with A transpose, almost the similar thing you can carry out with at A star actually. That is why this the definition and everything will go parallel hand in hand actually. And also remember, that if A star A transpose A is positive semi definite. Similarly, A star a is also positive semi definite guaranteed actually. Those are generalization of concepts from real to complex domain actually. Let us confine ourselves to A transpose A only as if that the singular value is nothing but Eigen values of A transpose A and square root of that actually.

So both A transpose A as well as A star A are positive semi definite for sure actually, that we have seen shown that before actually in the previous class. And if it is positive semi definite, their Eigen values are guaranteed to be greater than equal to 0; they are guaranteed to be either 0 or positive numbers actually. And hence, I will never land up with a complex value as a square root actually, because what is inside here is either 0 or a positive number. So, I have a choice of taking of plus or minus here, but I will simply consider the positive quantities and tell that that is my singular value actually. So, that is the singular value definition. So, as I told for singular value computation only positive square roots are neglected actually; negative square root you can you can neglect properly.

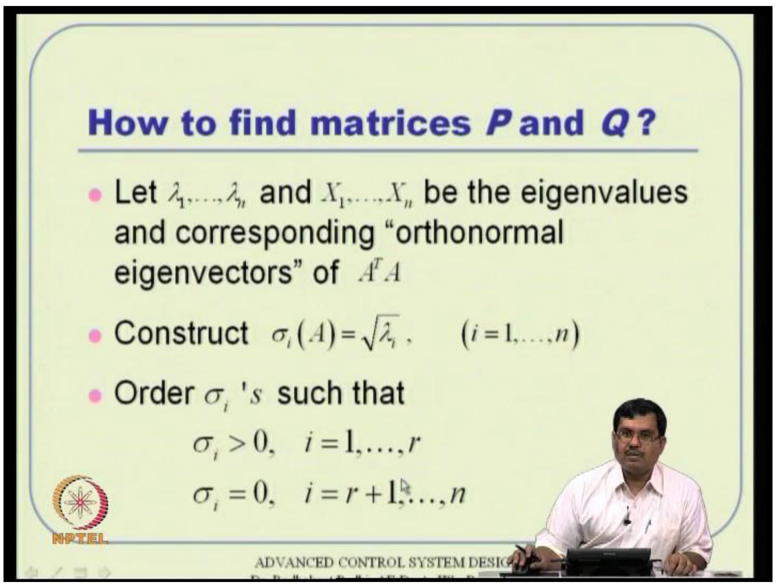
(Refer Slide Time: 20:03)



So, once you do this then this PDQ whatever you discuss, the D part of it will happen to be diagonal where in the diagonal you do not have 111 anymore, but you have sigma 1, sigma 2 up to sigma r now. If the matrix is of rank r, then the there exist orthogonal or orthogonal matrix is closely related to unitary matrix in complex domain. So, if you take PP transpose equal to identity then that tell PP star is also identity, that is that is in that sense it is called a unitary matrix in general and now let us not worry about that anyway.

So, there exist an orthogonal matrices P and Q such a way that A is equal to PDQ where D is nothing but this structure where the first diagram square matrix of r by r matrix the diagonal entries will happen to be singular values actually; just all other elements will happen to be 0.

(Refer Slide Time: 21:02)



Now, the question is how do you find this P and Q again. So remember, similarity transformation we had a procedure to find P and P matrix just took the Eigen vectors put them side by side and got the B matrix. Now, how do you find P and Q matrix both the matrix needs to be found out here actually.

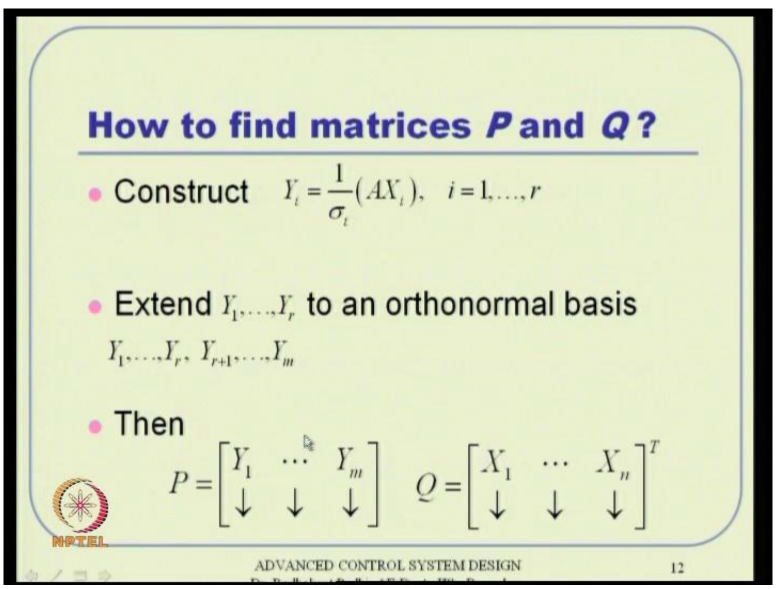
So the here is a procedure how do you find that. So, first we find out the Eigen values and Eigen vector pair for this matrix A transpose A remember that. This is this is see the definition wise it is actually Eigen values of A transpose A, it is not Eigen values of A only.

So, we have to find out the Eigen values Eigen value Eigen vectors are A transpose A that is the matrix Actually. Now, obviously by definition the singular values of A are nothing but the square roots of these Eigen values right, that is that is obvious from the definition actually.

So this Eigen these sigma's of the singular values I can find out simply from definition. Then, you can order these sigma's such that this remember A transpose A is going to be something like n dimensional vector actually. So you can have a bunch of 0s and a bunch of non 0 values. And the 0 values are not singular values by the way; by definition only positive numbers are singular values not 0s actually.

So you arrange that tell first r elements are strictly greater than 0, so those are the singular values and rest of the values are actually 0 0 0. So I will just order them that way, so that is just for convenience actually.

(Refer Slide Time: 22:35)



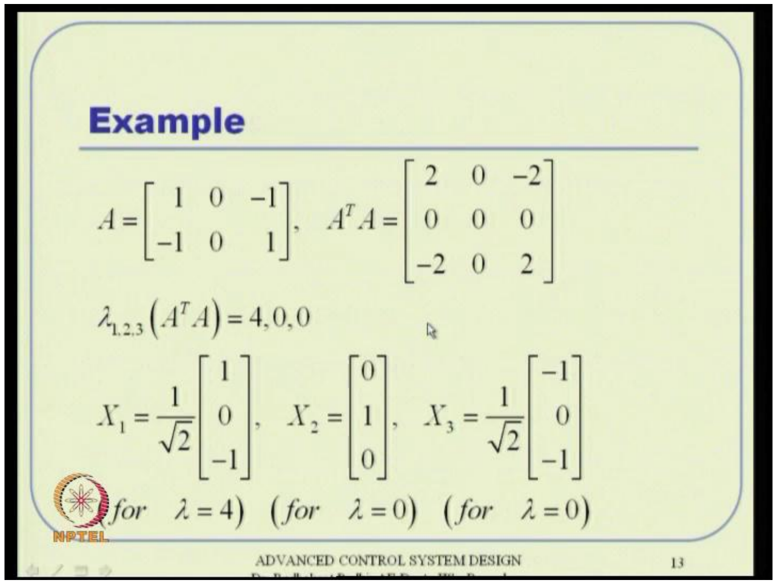
Now the question is how do you find this matrices P and Q still; I mean, that we did not discuss yet. So now, let us construct these Y vectors Y 1, Y 2 up to Y r this way; 1 by sigma i into a times X i. Remember X i is the Eigen vector for this matrix actually this A transpose A, so that is also we are having that by now.

So, Eigen each of the Eigen vectors if I just normalize that by dividing it by the corresponding singular value, then I will get this is remember A times that vector is also another vector and then I multiply that by 1 by sigma i, I will get another vector. So that vector I will construct Y 1, Y 2, Y 3 up to Y r and then tell that those Y 2, Y 1 to Y r Y r, I will put them together with further expansion actually. How do I expand that? So Y 1 to Y r is available in this series and I cannot do the similar algebra for rest of the thing, because rest of the things are 0 basically. So I cannot divide it by 1 by 0 and all that, that is not possible there.

So, but I still need an m dimensional vector actually right. So up to Y 1 to Y r I have already have. So, I have to expand the series or extend these basis vectors what do you call for further dimensions so that is possible actually. I will give an example, how do you how do you expand and all that actually next. So, I want to Y r is there so Y r plus 1 Y r plus 2 and all that we will we will construct that I mean it is possible to expand so that this these vectors will be orthonormal to this rest of the actually remember that. So this these columns whatever columns we need to put it later, all these columns needs to be orthonormal to each other, because P needs to be orthogonal remember that. So the column vectors has to be orthonormal to each other. Using that condition I will be able to expand this vector up to m dimension actually. Then, my P matrix is given by that and my Q matrix is given by that actually. Remember, x 1 to x n is already available by this by this Eigen value Eigen vector it is already available.

So, Q matrix is already there and P matrix is constructed that way and all that will be done for you by matlab command SVD inside that actually. That the routine whatever is available that will be doing this. It is also good to remember or understand what that command is doing actually. So this command is actually doing all sort of computation behind this algorithm actually.

(Refer Slide Time: 25:04)



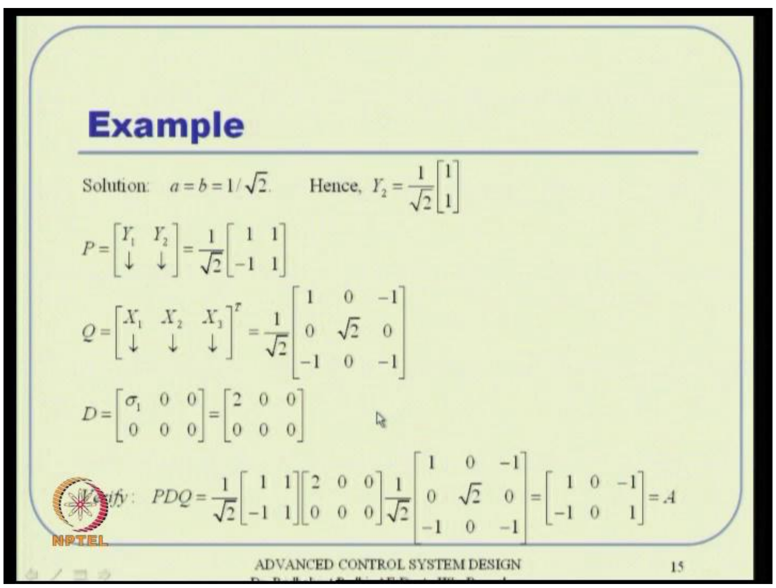
Let us see an example to make our ideas clear. We start with a 2 by 3 matrix; remember, this is not a square matrix; this is a 2 by 2 matrix simple 2 by 2 matrix. So first thing first step is to construct A transpose A. So I will construct A transpose A happens to be like that. Very quickly you can verify A transpose A if you remember it is 2 by 3, so A transpose is 3 by 2, 3 by 2 into 2 by 3 is actually 3 by 3. So hence, it you are getting 3 by 3 actually.

So remember, Eigen values of that Eigen value Eigen vector if you do that analysis; I mean, that characteristic equation and all that; lambda e minus this A transpose A then take determinant make it equal to 0, then you will find out that the Eigen values are nothing but 4 0 0 and here and obviously these are all kind of greater than 0 all this numbers, because A transpose A is guaranteed to be positive semi definite and you can see that from Eigen values also, either the Eigen values needs to be 0 or they are positive numbers and here it is actually.

So what is my singular value? My singular value is square root of 4; that means, my 2 is one of my singular value and that is it actually that is all I have and why is that? Remember, this rank of the A matrix is actually 1, if you see that the first row if I multiply by minus 1 I get the second row. So these two rows are not really linearly decoupled; they are linearly dependent actually. So that means, the rank of the matrix is 1 here, so I get only one square matrix; I mean one singular value actually.

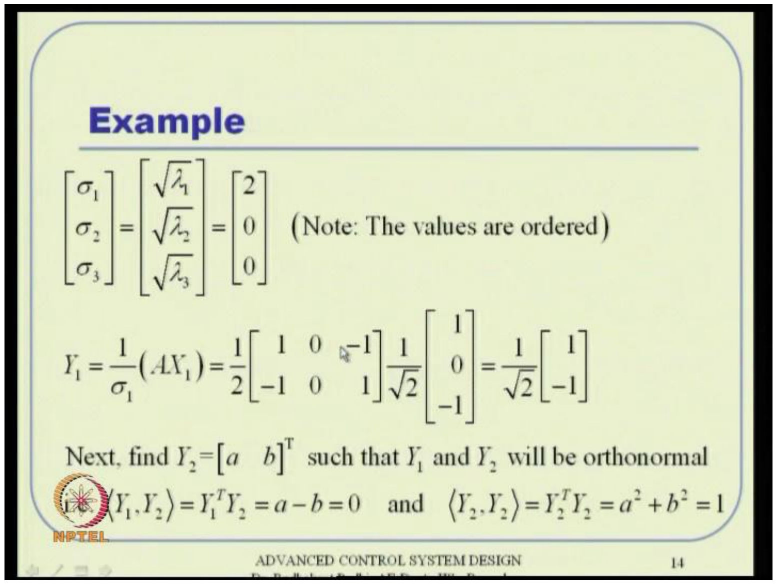
So, after I get this 4 0 0, I carry out this Eigen vector analysis and tell for lambda equal to 4 I have this Eigen vector and all these are normalize by the way we require that to be normalize, because these are all vectors of the column matrix Q. So this column matrix Q need to be kind if orthonormal to each other remember that. So each of the vector needs to be normalize also there. So I find out this Eigen vector normalize that, find out this Eigen vector normalize that and find out this Eigen vector normalize that and here is an example you can see that even though lambda is repeated 0 0, you will still be able to find out these two Eigen vectors separately; I mean, linearly independent to each other. And each of them are normalize; so each of them are actually orthonormal to each other.

(Refer Slide Time: 27:33)



So by finding out this x 1, x 2, x 3, I already get my Q matrix that I require later; x 1, x 2, X 3, if I simply put them together and take a transpose, this transpose is a essential by the way. You have to put them as either a row vectors in or you put them as column vectors and finally take a transpose, that is my that is my Q matrix actually. This is a this is a transpose here by the way. So, that is I found out x 1, x 2, X 3, then I put them together and then take a transpose so my Q matrix is ready.

(Refer Slide Time: 28:06)



Now, I go back and see what to do with the P matrix actually all right. So here is a 4 lambda 1; so my sigma 1 happens to be square root of lambda 1 that is 2. Sigma 2, sigma 3 are actually like 0 0; I mean, singular values are 0 0, so they will not play much of a role here actually. So, what you what you talk about that actually? Some by the way some definitions, some books will tell that 0 0s are not singular values also. The simply written like sigma 2 sigma, 3 does not mean they are singular values actually.

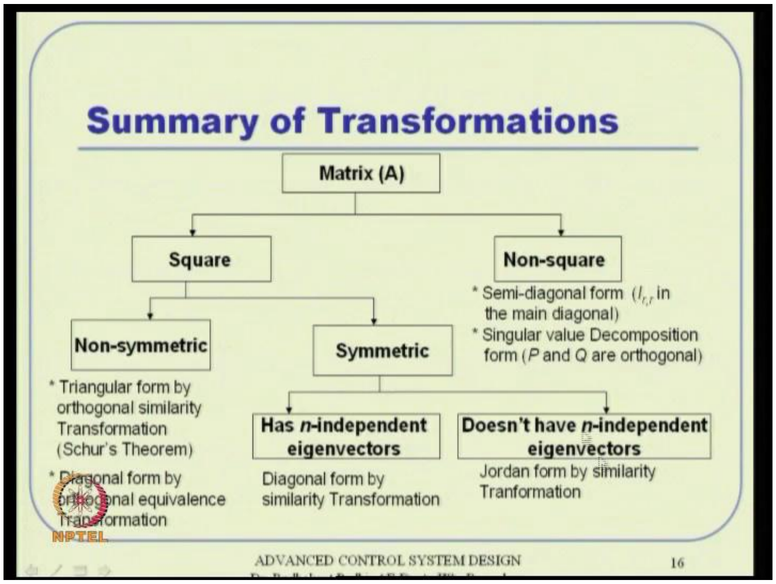
So anyway coming back to that, this Y 1 is actually 1 by sigma 1 times AX 1 by that procedure. Remember that procedure you have to carry out now this 1 right. So, sigma 1 is non 0, so we you can get Y 1 like that. So Y 1 is straight forward 1 by sigma 1 to A times AX 1. So, I will carry out this exercise; I will multiply a times x 1 then take out 1 by sigma 1. So, this is A matrix; this is 1 by sigma 1; this is A times x 1 and finally I will end up with this vector actually 1 by square root of 1 minus 1. So this is all about Y 1.

So, what about Y 2 now? We have to find extend this Y 1 to Y 2 in such a way that Y 2 will be orthonormal to Y 1 that is the condition actually. So, let me start with Y 2 as ab; a b column wise, so I put A transpose here actually; that means, a and b are column wise next to each other actually. So I find Y 2 in general such that Y 1 and Y 2 will be orthonormal. So that means what? The inner product between them is going to be 0 and the inner product between Y 2 and Y 2 has to be 1, that is the condition actually.

If I Y 1 and Y 2 are inner product and if I carry out that is a minus b which is equal to 0 and inner product of Y 2 and Y 2 that is Y 2 Y 2 transpose Y 2 and things like that. That is Y a square plus b square is equal to 1 actually. So, I have to satisfy this equation; that means, a minus b has to be equal to 0 and a square plus b square has to should be equal to 1; and a minus b equal to 0 means a is equal to b. Now you put it back here, then it tells me that 2 a square equal to 1, and hence a is nothing but 1 by square root 2 basically. So obviously, this is also equal to that, so a and b are equal to 1 by square root 2; that means, Y 2 is nothing but 1 by square root of 2 1 1 actually.

So, I had Y 1 like this and I extended that Y 1 to Y 2 such that Y 2 is orthonormal to Y 1 and in that process I got Y 2 actually. Now, I have P matrix is Y 1, Y 2 which is given by like this and Q matrix already I have actually. Now, if I carry out that algebra PDQ sort of thing, then d turns out to be only sigma 1 and everything else is 0 actually here; sigma 1 is 2 and you can clearly verify that PDQ, if I do this algebra this is my P matrix; this is my D matrix what I am proposing here and this is my Q matrix and if you do this algebra, naturally it turns out to; I mean, very nicely it will turn out to be a and I mean in my suggestion for your if you do some exercise in exam or other then; I mean, just verify this equation one time at the end if you that happens to be there then you are safe actually.

(Refer Slide Time: 31:29)



So, what is summary of transformation? If you see this the many things that we talked Actually. So, A matrix in general can either be square or non-square right. If you have a non-square matrix, you can actually you can reduce it to semi-diagonal form only; semi-diagonal form is that form that we discussed about this at the beginning itself right. We discussed about this form I being there and the main diagonal actually.

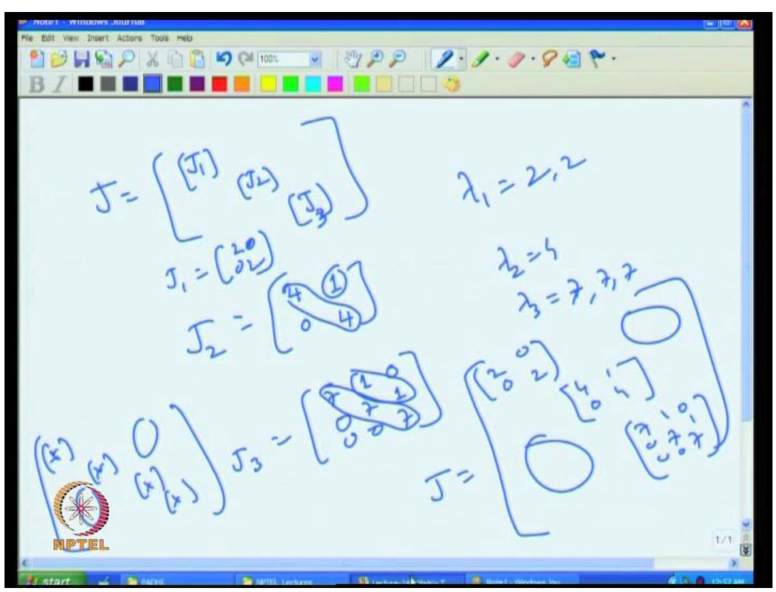
So either you can go that I r r in the main diagonal, that is a semi-diagonal form or you can question that and tell hold on a second, we can actually go ahead and carry out singular value decomposition. And this is actually a lot of meaning, because many of the numerical methods by the way, the conditioning and other thing the whether your numerical algorithm is well conditioned or ill conditioned things like, that they also come from something called condition number. And condition number is nothing but maximum singular value divided by minimum singular value.

So singular value is not just about matrix transformation alone. They there lot of usage for numerical methods as well actually, so that is about non-square. So, if you have a non-square matrix, you have only this or this singular value decomposition primarily as you as your; I mean, some sort of a weapon for reducing the matrix actually. Now what about if you have a square matrix. Now that will arise two different cases, either I have a symmetric matrix or I have a non-symmetric matrix.

Now, if I have a non symmetric matrix, I can directly go ahead and apply this cross theorem and tell well I can reduce, I can restrict my P matrix to be orthogonal, then I will end up with triangular form by orthogonal similarity transformation basically instead of the P inverse AP; P has to be orthogonal as well. Then I will end up with triangular form or I can actually go ahead and do this orthogonal equivalence transformation like what I do here for non square matrix, I will also do the same thing here and I will end up with a still a diagonal matrix actually right. But in the diagonal, I need not get this one this… If I do that way in the in the diagonal element I will get similar values actually.

Now, what if the matrix is square and symmetric? It is square as well as symmetric matrix. In that situation it can have n linearly independent Eigen vectors or it need not have that. And if it has n linearly independent Eigen vectors, I can reduce it to diagonal form by similarity transformation where the diagonal elements are nothing but Eigen values or if that is not possible, then I will end up with Jordan form by similarity transformation. And what is Jordan form? Jordan form consist of… Let us talk a little bit about that. What is Jordan form?

(Refer Slide Time: 34:32)



The Jordan form is something like J which will which will talk about block Jordan form actually; J 1, J 2 block Jordan block diagonal matrix where these are this J 1, J 2, J 3, like that are actually Jordan blocks actually. And what is that? Let us say J 2 is transfer lambda 2 which is repeated twice, then if it does not have two different Eigen vectors, then I will end up with this lambda 2 is let us say 1 actually, so then you will end up with 1; 1 is not a good number probably take let us say 4; lambda 2 is actually 4, then corresponding J 2 will have 4 4, then it will have 1 here and then 0 here not 0 actually there. So this particular thing what you see there that at this particular thing is not 0 actually, that is that is 1 here actually.

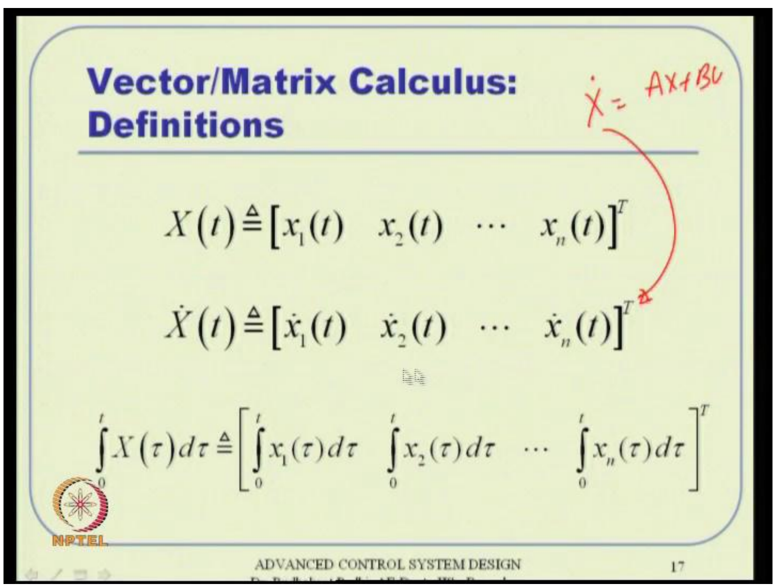
Now, if you have let us say J 3 lambda 3 is lambda, lambda 3 is equal to let us say 7 or something, but it is repeated thrice let us say 7 7 7 sort of thing, then corresponding J 3 will have 7 7 7 here and obviously these are all 0 0, the below side of it. The above side it will have 1 here and 1 here and then 0 here. So you have this 1 is coming up just above the diagonal sorry this is not true actually. So you have 0 here and you have this this this is the diagonal thing. And similarly, this is diagonal 7 7 7 and then there will be a 1 actually. And remember, if J 1 this lambda 1 it happens to be 2 and then it is just I mean kind of it is repeated still, but you have sorry it is it is repeated, but you have linearly independent Eigen vector for that, then the corresponding J 1 will have well the corresponding J 1 will have simply 2 2, this 0 this 0 will happen here and still it is a diagonal matrix actually. So, if you put them together your J matrix will turn out to be a very big matrix and that is first 1 is 2 2 anyway so I will not worry about that, that is my J 1.

That is the second block thing will be having like 4 4 1 0 and the third element will have 3 that is like 7 7 7, then 1 1 here then 0 and everywhere else it is 0 anyway, so that is that is Jordan form that you are talking about actually. So, that is the that is the in general you will be able to do that actually all right . So this is a summary here; what it tells if I have n independent Eigen vectors, I will end up with a diagonal matrix or if I if I do not have an independent Eigen vectors, I can land up with a Jordan form in general and this is actually generalization form over this actually.

And if you really want a Jordan form, then that is where your generalize Eigen vectors will lead you to Jordan form actually. If you have regular Eigen vectors, it will lead you to diagonal form and diagonal form is also like a special Jordan form you can say that way, because if you have a if you have a diagonal matrix for say something here something here something here and each of that is also a Jordan block actually. So, a diagonal matrix is also like a Jordan matrix in general actually, so you can say that way.

So, this is all about transformations actually; summary of transformation talks about that. First you have to see whether you have a square matrix or you have a non-square matrix. If you have a non-square matrix options are limited, we are just talking about a singular value decomposition. If you have a square matrix, you can have a non-symmetric square matrix or a symmetric square matrix. A non symmetric square matrix again the choices are not very high. But if you have a symmetric square matrix, you can see the beauty actually. You can have like an independent Eigen vectors, and then you can does not have an independent Eigenvectors, and then you can have diagonal form, non-diagonal form and Jordan form like that actually you can talk.

(Refer Slide Time: 38:50)

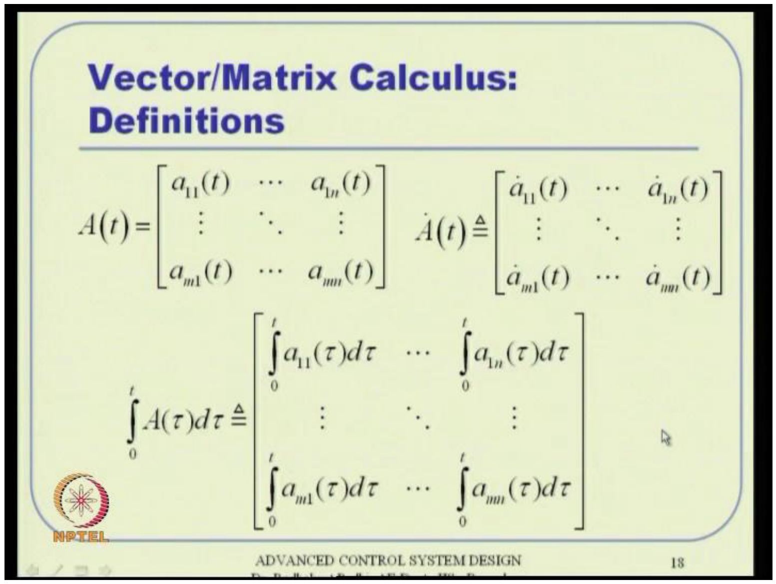


Anyway, so these are all the transformation actually. Now let us talk about different concept all together and that is what we will need in control theory also many times. So, here is a concept to which we call as vector matrix calculus actually. So, what we are doing here? We are talking about taking calculus further into the vector matrix domain actually. We know what is differentiation, we know what is integration in a scalar sense actually. Now in a vector sense what is happen; in a matrix sense what will happen.

And remember, most of the time we do talk about I mean without even bothering for that we talk about X dot equal to AX plus BU many times we see that actually. So, what is X dot after all? X dot is nothing but what X dot you see here is something like that what you what you define here. So that is nothing but x 1 dot, x 2 dot, x 3 dot. So, if I have x's like this x 1, x 2, x 3, which are all time varying let us say and time is just a free variable; you can just assume that is a some independent variable and the derivative is with respect to that actually. So, if it is t here which is very common in dynamical systems anyway, then we talk about X dot or dx by dt actually.

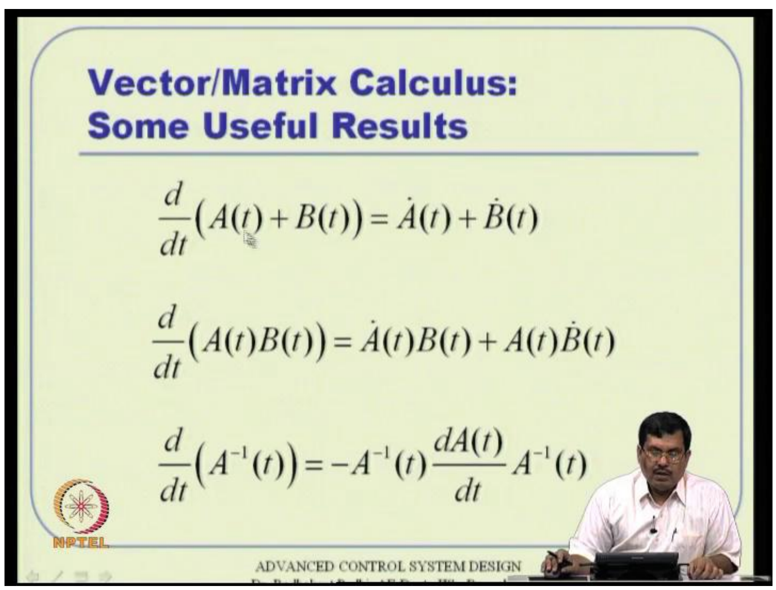
So by definition what you do here is, if you have x 1, x 2, x up to x n, then by definition X dot is x 1 dot, x 2 dot, x 3 dot, like x n dot simply by definition. Remember, these are all column vectors I will just put them row by row and then made a transpose actually here. These are all column vectors actually that is a default, default notation is column vectors in control theory. So, if you similarly if you talk about integration of a vector, then all that you have to do is element by element integration. So, differentiation is element by element by element differentiation and integration is element by element integration, that is all you do actually.

(Refer Slide Time: 40:41)



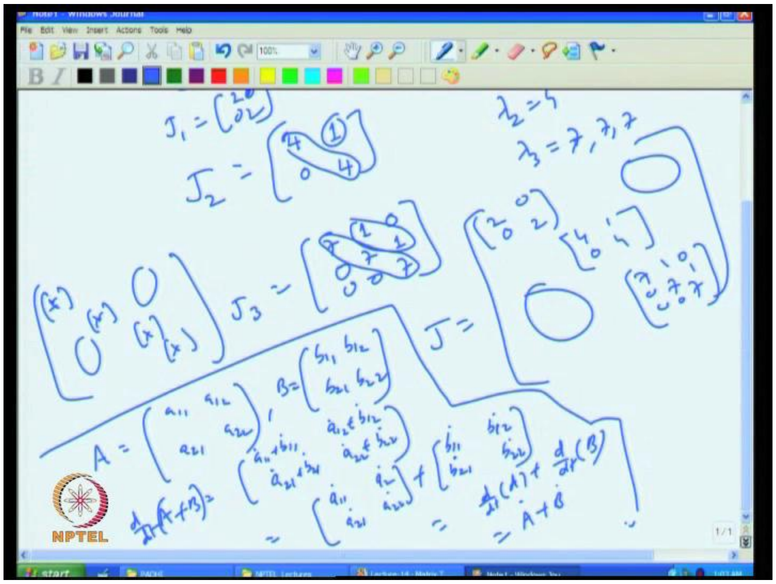
Similarly, if you have a matrix which is time varying, then you can talk about differentiation of that A dot is nothing but element by element differentiation all over and integration is element by element integration all over.

(Refer Slide Time: 41:00)



And then it turns out some of the useful results which will tell what will happen if I have d by dt of A plus B, then it is very clear that it will have A dot plus B dot simply from definition actually. If I have A dot something like this; A is something like this, B is something like that, then I can d by dt of B let us probably do that very quickly that is rather easy actually.

(Refer Slide Time: 41:34)

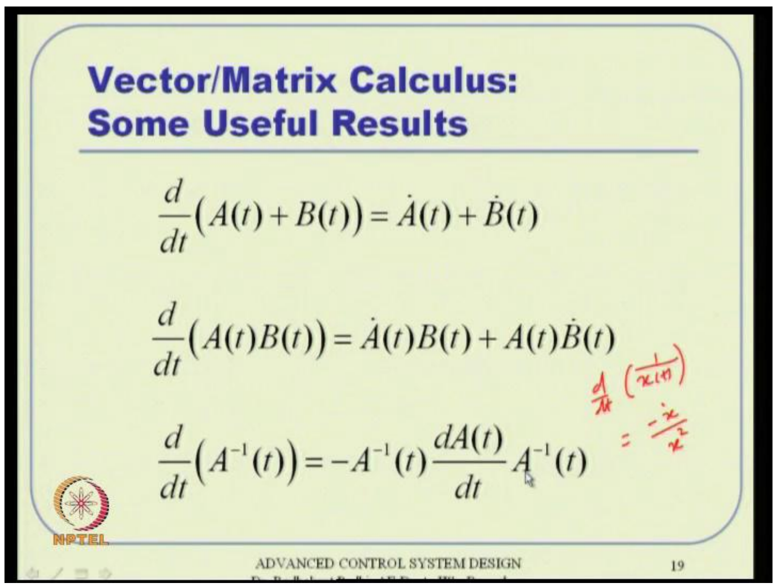


So, I have A matrix something like a 1 1, a 1 2 let me take a 2 by 2 for example, a 2 1, a 2 2 and similarly, B matrix which I have b 1 1 and then b 1 2 and b 2 1, b 2 2, then A plus B is nothing these are time varying elements remember that, then A plus B is nothing but a 1 1 plus b 1 1, a 1 2 plus b 1 2, a 2 1 plus b 2 1 and a 2 2 plus b 2 2.

So, if I talk dot of that dot of that, then this is actually d by dt of this quantity A plus B. So this is nothing but d by dt of the whole quantity I can divide into that, so this is a 1 1 dot plus b 1 1 dot, then a 2 1 dot plus b 2 1 dot and then a 2 1 like that actually. All that thing I can do, then I can decompose that and tell I can take out this and tell this is a 1 1 dot a 1 2 dot here; a 2 1 dot a 2 2 dot plus this plus b 1 1 dot b 1 2 dot b 2 1 dot and b 2 2 dot. So then this is nothing but by definition I can take out this derivative 1 outside again and tell this is nothing but d by dt of A matrix plus d by dt of B matrix, so this is nothing but A dot plus B dot.

So simply from definition, we will be able to show that; I mean A plus B dot rather this is actually what we started with is a nothing but this this this quantity is actually d by dt of A plus B. So, this d by dt went inside this d by dt went inside of that, then you talk about derivatives everywhere, then you separate it out then tell this is nothing but A dot plus B dot actually. So simply from definition, we will be able to show this actually all right. Now, go back to that and then tell for the these are all simply definitions anyway.

(Refer Slide Time: 43:49)



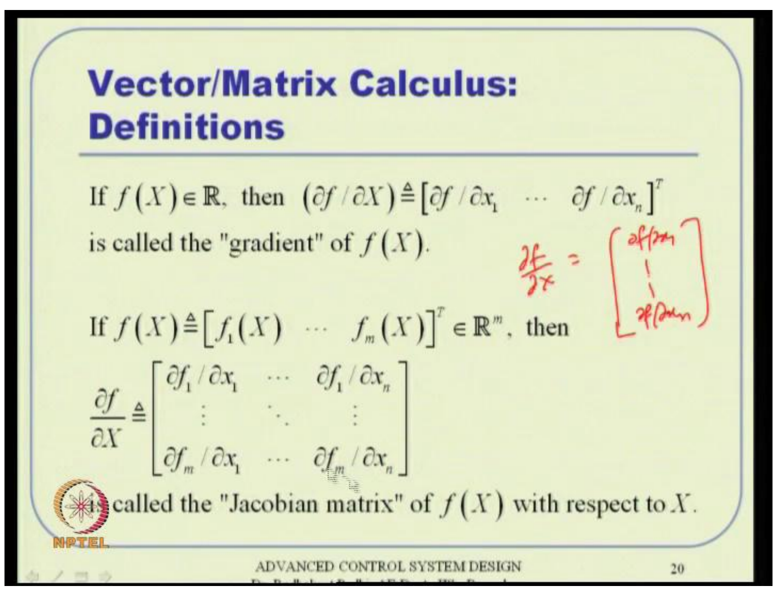
So, these are some of the nice results which will tell me that A B A plus d by dt of A plus B is nothing but A dot plus B dot actually. So, we use that many times implicitly also I mean is true. Now, what if there is a multiplication? Obviously, these matrices have to come; I mean, this multiplication should be well defined actually. So you know how this matrix multiplications are there. So, if I use m by n then this has to be n by P somewhere, so the multiplication has to be like m by P after all actually.

So, this particular thing in a scalar sense you can already see that A and B are scalar, then that is the result. So, it is also in general true for vector; I mean, matrix multiplication in general, and the only thing that you need to keep it in mind is that the sequence of multiplication should be preserved. You cannot write A dot into B is not equal to B into A dot that is not true in general. So, the sequence of operation has to be preserved actually, then this result is also true. Again you can do that longhand multiplication in general; you take a 1 1, a 1 2, a 1 3 like that up to m by n, then correspondingly you take B, then you apply the definition, then it will go element by element derivative. Then element by element it will pop up that way, then you take out and ultimately you will combine them like this actually. It is not very difficult to show from simply definitions and elements actually.

Now here is an interesting thing which tells me that derivative of the inverse of a matrix A is nothing but negative of A inverse d A by dt into A inverse actually. So, it is very close to what I take it. Suppose, I have a scalar quantity then the d by dt of this is nothing but x inverse A; that means, 1 by X t let us say if it is a scalar quantity, then we know for sure that this is actually x minus of x dot divided by x square actually all are time varying by the way.

And similar thing you see here, but I cannot see it as the 1 by terms and all I cannot write it there. So, what I what it turns out is very close to what we know for scalar actually. This is for scalar and this is actually for matrix calculation actually and we can show that the proof is not very difficult actually the further results.

(Refer Slide Time: 46:13)



Now, we will require some sort of kind of a Jacobin matrix and things like that also. If f(X) is this now it is a function of state which is actually X is a function of time we know that.

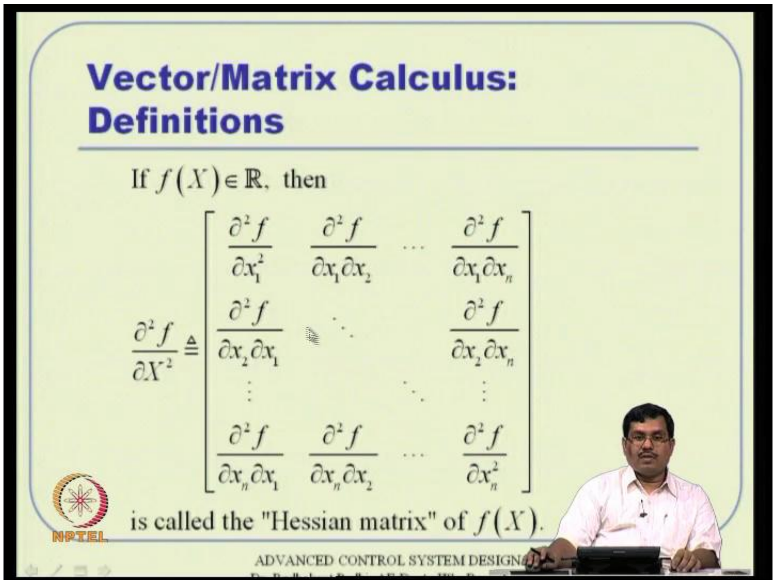
But in general, we will not worry the time dependence part of it. Let us simply tell that I have a function of vector; that means, the f is a scalar, but X is a vector and it is possible right; one example is x 1 square plus x 2 square plus x 3 square; I mean, that is if you have X is x 1, x 2, x 3, then the function is what I am having x 1 square plus x 2 square plus x 3 square is actually a scalar quantity, but X is a vector quantity actually. So something similar situation I am taking actually. So X is a vector, but the function of that vector is a scalar quantity.

Then, what you define is a Jacobin vector; I mean, this gradient vector rather here. The gradient vector is defined as something like this. Remember, this again this is a column vector actually, this is A transpose out here. So, how do I define? I just take the partial derivatives del f by del x 1 del f by del x 2 like that, I simply put them together down by down actually. So, what I mean here is del f by del X I define as del f by del x 1 del f by del x 2 like that del f by del x n simply by definition, it is a column vector.

Sometime people define it as row vector in some books, but throughout our course we will follow that as a column vector actually; whether there are small advantages of X defining whatever way, we you have to be just consistent with whatever definition you have actually. So in our throughout our course, we will take it as a column vector and this is nothing but the gradient vector also del f by del X is nothing but a gradient vector.

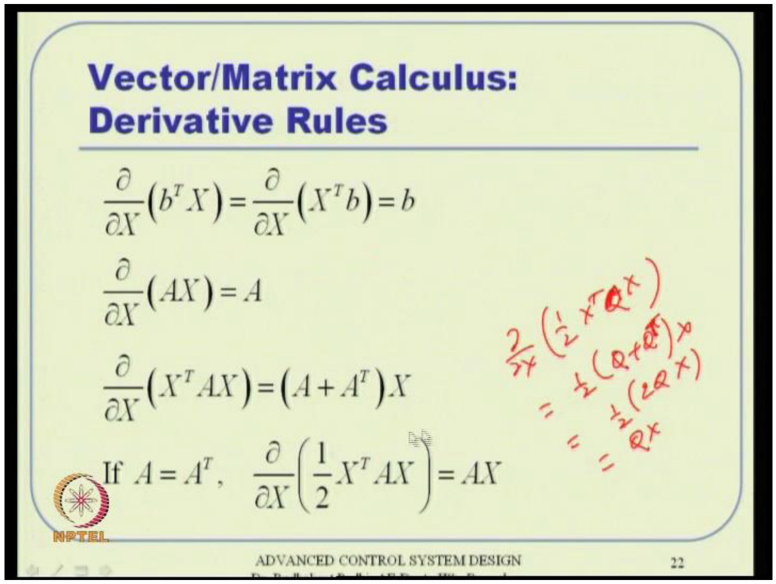
Now in the next case, what if I have a vector function; that means, X is a vector and f is also a vector; they need not be have some dimension by the way right. X belongs to R n, whereas f belongs to R m, so I have a I have a vector function of a vector actually, then I can I can still take the first partial derivatives and put them together in a matrix format and this matrix is called Jacobin matrix actually. So, if you have a scalar function, this vector is result in a vector which is called a gradient vector and if is a vector function, it will result in a matrix which is called a Jacobin matrix.

(Refer Slide Time: 48:46)



Then, the next concept what if I have a I have a scalar function of a vector, but I take about take up second partial derivative del square f by del X square. How do I define that? And I can still go ahead and define that this way what you see here. And also remember, that del square f by del x 1 del x 2 is also equal to del square f by del x 2 del x 1 that is a standard result from calculus actually; that means, this Hessian matrix is guaranteed to be a symmetric matrix. Now, this lecture and previous lecture we have seen very neat properties about symmetric matrix. Many of this things will be exploited later in especially these matrices are kind of they will appear in optimization theory. So many times it will end up with hessian matrix and things like that for sufficiency conditions, and sufficiency conditions is nice to see from second derivative and all second derivative happens to be hessian matrix which is already a symmetric matrix. So these are all nice properties which will be exploited there actually.

(Refer Slide Time: 49:52)



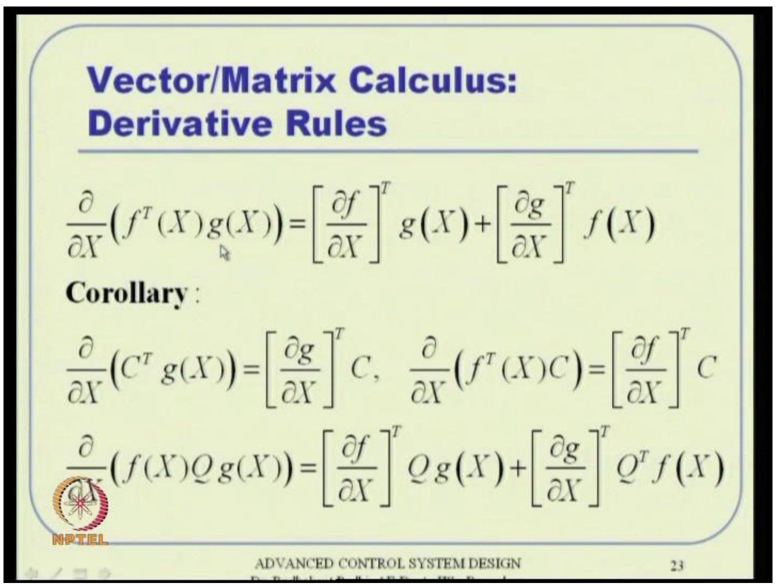
Now, further results some of the results are very useful later. Suppose, I have a constant b constant vector b and there is X actually. I talk about taking partial derivative of X with respect to b transpose X and remember b transpose X is a scalar and X transpose b is also a scalar. If I take partial derivative it happens to be simply b, and you can just do it by longhand algebra, and so that very easily actually. You can simply expand that b transpose is X is nothing but b 1 x 1 plus b 2 x 2 plus b 3 x 3 like that and if you take out this partial derivative by definition, it will simply pop up as B matrix actually. It is not difficult to show at all actually.

And then, if you do that second result tells us that, if I take AX and then take partial derivative del X by del by del X of X it happens to be A. Again you can do that by longhand algebra you can show that actually. Then, I take what if what if I have a matrix something like this X transpose AX, then I will I will use that property wherever property I have seen here something similar to that and I can generate, I can really land up with this result actually. X transpose AX if I just take del by del X, it will result in A plus A transpose times X. And as a corollary, if I have a symmetric matrix and I talk about del by del X of half of X transpose AX where A is a symmetric matrix, then what let me generalize that a different matrix just to avoid confusion and all and that is also necessarily that this is a Q matrix let us say.

Then, what is that? This is actually half of Q plus Q transpose right Q plus Q transpose times X and then Q transpose is equal to Q basically right. So, what you get? Half of 2 Q X basically that is nothing but Q times X. So, if you have a symmetric matrix to begin with Q is a symmetric matrix, then what you what results in just Q X provided there is a half term already there. So, these are the reasons why you see now then LQR theory a term something like that linear quadratic regulator theory when you talk about that, we will see this term as part of the cross function actually.

And, we will also need a partial derivative with respect to X for what is known as equation and things like that. So there these algebras should be useful actually what you see here. That is what I am telling here if a is a is equal to A transpose, then this is the result actually.

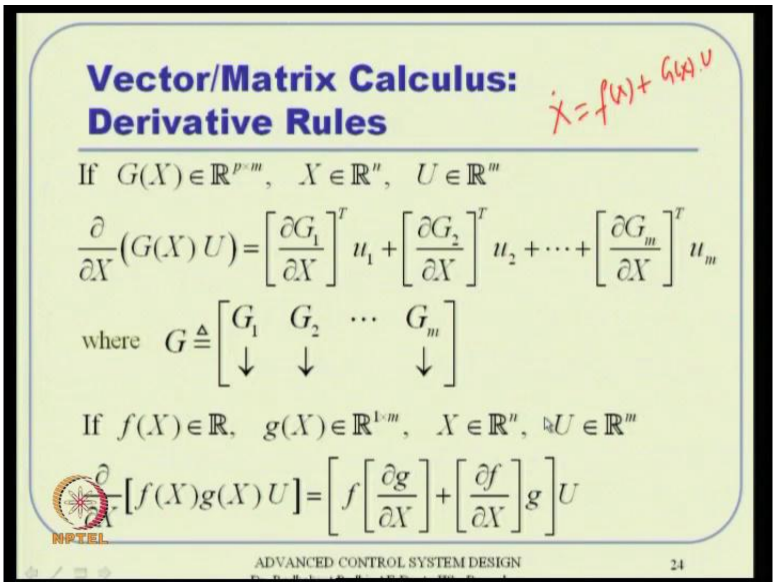
(Refer Slide Time: 52:35)



Now for the results what is known as chain rules and all that actually. So, this kind of algebra we will we will encounter in optimal control theory primarily and some of this non-linear control theory as well. So, what if what if talk about f transpose g now. Now, f is a function by of X; g is a function of X; they are dimensionally compatible and f is a vector, g is a vector. So f transpose g is a scalar and then you can still talk about taking partial derivative with respect to X and this happens to be like this actually.

And these are all not very difficult to show at all actually, just by longhand algebra you will be able to derive that actually. Now as a corollary, now one of the function happens to be simply a constant; that means, f of f transpose X is f is nothing but C actually, just a constant vector, then you will land up with something like this. This is just from this result you can derive those corollaries actually, and if you have a have a matrix Q in between by f transpose Q g X, then it will also happen to be something similar to that actually. You can you can take Q times g X is some g 1 of X or something and apply that actually what you have here, and then you will land up with some formula like that.

(Refer Slide Time: 53:55)

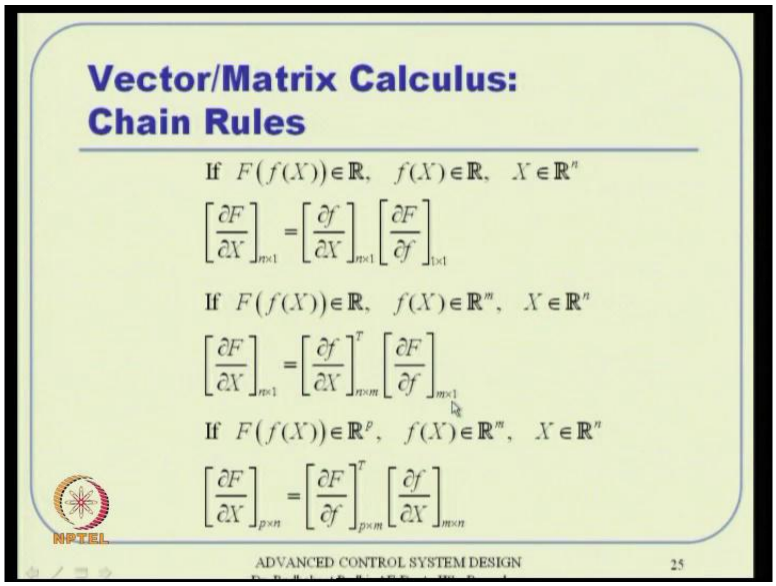


Further generalization, if you have a have a matrix now matrix function of X and there is a vector u also, then the partial derivatives can also be derived and why do I talk about all this actually? Remember, we have a special class of functions systems of what is called as control non-linear systems where you talk about f(X) plus G(X) times U; G(X) is a matrix actually by the way.

If you have a I mean these are special class of systems what we will deal with some time later; f(X) is a vector and G(X) is a matrix actually. So if you have a control system non-linear system that way, then some of this algebra will encounter actually on the way. So, let us get ready for those things and then tell these are the results that we will need later actually. And obviously, I do not want I mean there is not require to be remembered actually, I mean these formulas and all we need do not have to remember.

However, we have to know all these results because many of the results will be useful I mean you should you should not be afraid about these results actually.

(Refer Slide Time: 55:06)



Now, there is a further question function of functions what will happen actually. So, if I have a function of function, then the results are like either I can have this small function as a scalar and a big function as a scalar as well; or the small function is a vector, the big function is a scalar; or the small function is a vector and big function is also a vector; that means function of a function.

So, first function is scalar or vector; second function can be scalar or vector. Then, I talk about derivatives of that final function, then obviously what comes to mind is chain rule actually right. Function of a function if you take derivative that is chain rule. So the generalization of chain rule to vector matrix thing is also available, and these are dimensional compare to compatibility sense it will be maintained, but the results are fairly similar to what you already know actually. Just have to be careful about the dimension compatibility and formally you can show that also the derivations and all are possible to show that actually.

So, these are all vector matrix calculus actually. So with that I will probably stop here and so message here is we will now; I mean, this these algebras are possible and we will use that algebra wherever it is necessary actually. So, with that as that last class and this class we have seen many of the concepts from matrix theory which are helpful for dynamical system analysis as well as control design actually. So, we will I will come back to this control design concepts in our subsequent classes and with that message, I will I will stop this actually, thanks a lot.