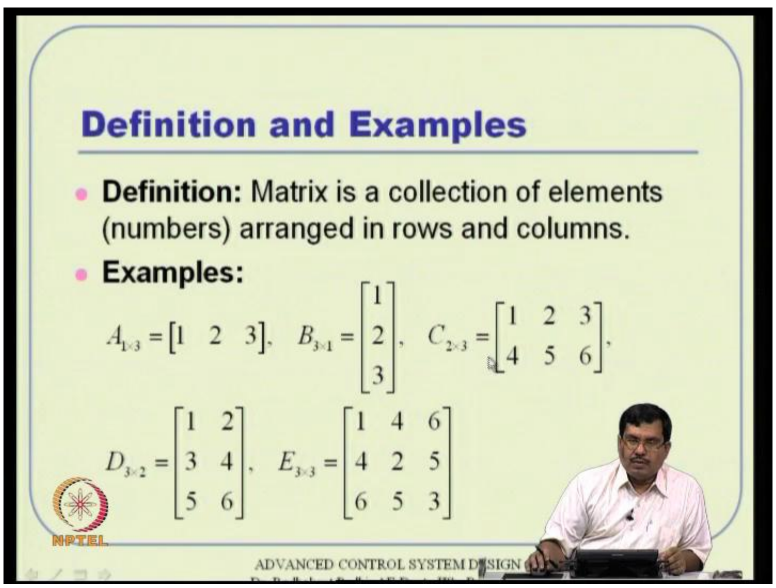
Advanced Control System Design
Prof. Radhakant Padhi
Department of Aerospace Engineering
Indian Institute of Science, Bangalore

Lecture No. # 12
Review of Matrix Theory – I

Good morning students, we have so far we have covered many topics, and it is probably very clear to you that we need lot of matrix theory for our course through out. So, control theory, modern control theory especially demands lot of matrix theory concepts. And hence I thought we will review some of these matrix theory concepts in a couple of lectures actually. Now, we will proceed with this, this basic concept in this lecture followed by some of this slightly advanced concept in the next one, and then move on from there actually.

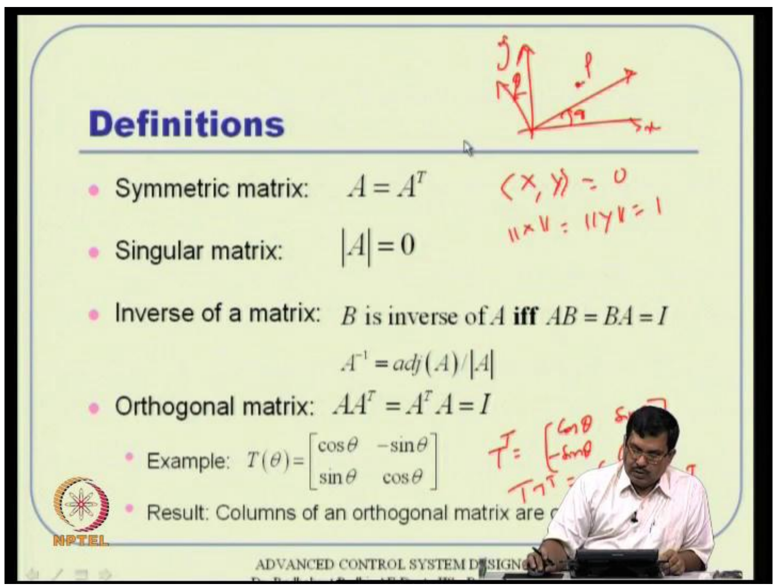
(Refer Slide Time: 00:56)



So, what is basic definition for a matrix? It is essentially nothing but a collection of elements and typically numbers arranged in rows and columns as very standard, and I think all of you probably know this. So, if it is a various ways you can arrange the numbers, you can either you can arrange in just a row or just a column or a combination of a rows and column. So, if it is just a row, it is a matrix also. Any single element is also a matrix - 1 by 1 matrix, but here we have something like a rho vector, what is and this is what is called a column vector also. And then this is like a 2 by 3 matrix, and this is a 3 by 2 matrix, and there is a 3 by 3 matrix like that actually.

So, depending on how many rows and columns you have, you can also give something like a suffix here, I mean this 1 by 3 sort of thing here and then 3 by 3 here and all. But these are not required, because it is very implicit from matrix dimension actually. So, implicitly we just take it for granted actually, so we may not be able to write this subscript all the time actually, need not do that.

(Refer Slide Time: 02:06)



And, then some further definition we call a symmetric matrix, a matrix is symmetric if it satisfies A is equal to A transpose. That is if you take a transpose of the matrix, what is transpose by the way? This if you exchange rows to columns and then columns to rows, then it becomes like a transpose. For example, if you just see here and then if it is, if we really takes C transpose, then 1 2 3 will be down and then 4 5 6 will be next to each other thing, next to that. So, that is what is called let us say if you, I mean if it is C transpose for example, then this will be something like 1 2 3 here and then 4 5 6 there, that is the transpose.

Now, obviously we cannot talk A equal to A transpose, unless A is a square matrix actually. So, we need A to be a square matrix, so before we talk about A equal to A transpose. But all square matrix need not be symmetric either; it has to satisfy this property for a matrix to be symmetric. And, especially if you see this particular matrix and obviously it is not, I mean it is actually a symmetric matrix where you see this element is coming here, 1 2 element is actually 2 1 element and this element is also here and this element is also there. So, A is actually a symmetric matrix that way.

Now, there is a some concept called singular matrix and singularity means the determinant has to be 0. And, these concepts come from inverse definition also, because inverse is defined as adjoint of A divided by determinant of A. So, if the determinant is 0 then A inverse tends to be, it tends to go good infinity actually. So, singular matrix is A matrix for which the determinant is 0. Now, by what is inverse of A matrix, that is a next concept. The definition wise B matrix B, is an inverse of matrix A if, this means if and only if A times B is equal to B times A is equal to identity. So, if that happens, then B and A are inverse of each other actually.

And, inverse of A is typically given by this formula, but this also as a definition actually A B equal to B A equal to I is the definition. And, A inverse happens to be adjoint of A by determinant of A, I am sure very sure you know what is adjoint of A, you have to find out the minors cofactor and then make some transpose and all that actually. So, I think that is known to all of you. Now, the next concept is what is called is orthogonal matrix. An orthogonal matrix is a matrix for which you satisfy this property, A times A transpose is equal to A transpose times A is equal to I. What is the duty of that? Now, correlate that to this inverse matrix definition, you will find that A inverse is nothing but A transpose. It is a very beautiful property; we do not have to do someone, so much of operation unnecessarily. This orthogonal matrix inverse happens to be just the transpose actually.

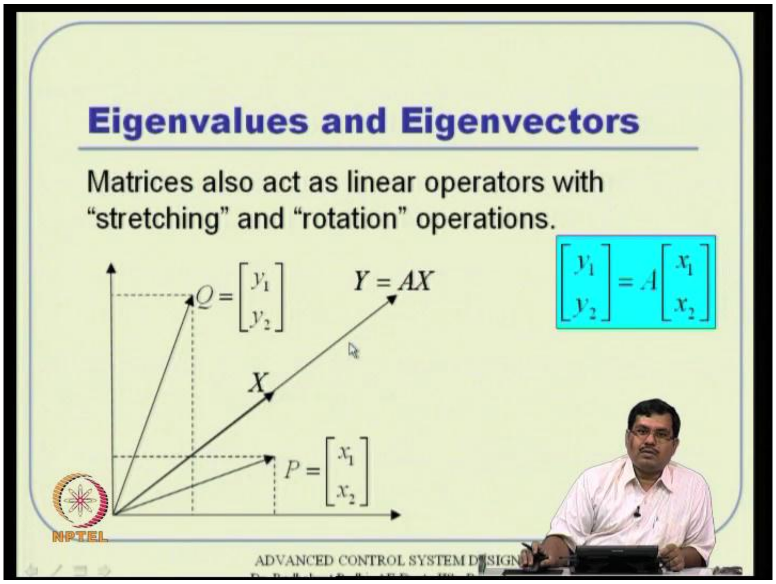
And, then what is an example for that? If you have a, this kind of A matrix cos theta cos theta and minus sin theta sin theta out here, then you can very clearly see that if I just take a transpose of that, very quickly probably you can do that here. That let us see T transpose is a something like cos theta cos theta out here, and then it is like sin theta minus sin theta out here. That is row going to column and column going row basically. Then, if you multiply this together this A matrix, this T and T transpose if you multiply together, then you will get cos theta times cos theta that is cos square theta. And, then it will get that cos theta sin theta and minus sin theta cos theta will cancel out, so what will happen here is like T times, sorry for that, the T times T transpose if you do that, it will happen T times T transpose something like one here one here and 0 here 0 here, that is nothing but identity actually. So, that is why this matrix is actually a orthogonal matrix out here.

And, also you can see that this particular example it is very clear. That if a matrix is orthogonal, then the columns are orthonormal to each other. The columns are vectors and it is orthonormal, orthonormal is what x and y vectors, vectors x and y are orthonormal provided the inner product between them is actually 0. And, then magnitude of x and magnitude of y is both 1 actually. And, you can see that because cos square theta and if you have this property, then cos square theta plus sin square theta equal to 1 and it is also happens to be the same thing for column 2. And then if you take x, I mean inner product of that cos theta into negative sin theta, that is minus sin theta cos theta plus sin theta cos theta that happens to be 0. So, inner product is 0 and the magnitude of each column is equal to 1 actually.

So, magnitude or inner product mean 0, they are orthogonal to each other and magnitude being 1 in addition to the orthogonality, we call that is orthonormal vectors actually. So, if a matrix is orthogonal, then the columns are suppose, I mean they are orthonormal to each other. So, that is a property that is we, that we should keep in mind actually. And, by the way this matrix that you see here is also a very neat matrix; it is what is called as rotational matrix actually. So, if you have for example, if you have a coordinate frame something like this and they have a point here, point p let us say and you have some coordinate in this X and Y, now the question is suppose I rotate this coordinate frame by angle theta, this coordinate by angle theta, then what is the same coordinate in this new coordinate frame? So, that property is related to this matrix that you see here.

In other words the same coordinate in a, in the new coordinate frame will be this matrix times double coordinate actually. So, we will get the new coordinate, it is rather quite easy to derive that also actually.

(Refer Slide Time: 08:49)

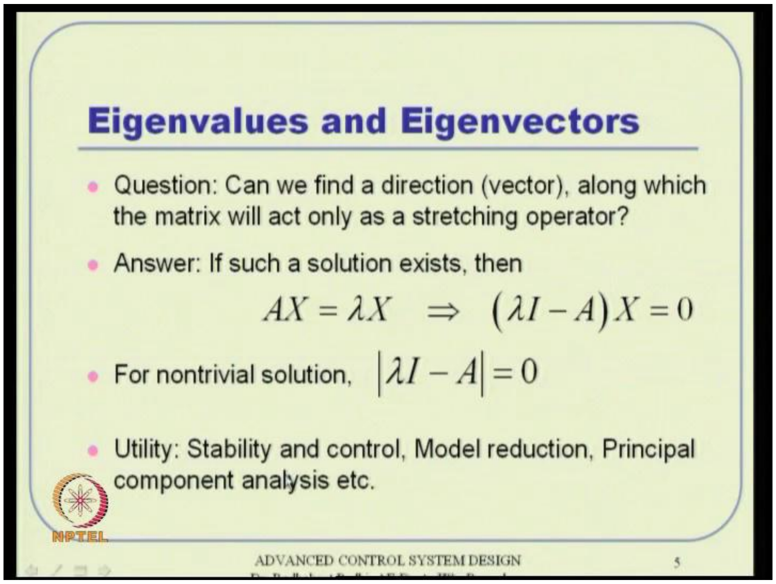


So, we will proceed further and then concept, the next concept that we require heavily in control theory, is the concept of Eigenvalues and Eigenvectors. Before we know this in little more detail, we also realize that matrices are also linear operators, any matrix where matrix numbers what you arrange them; they also serve as linear operator. And, that has a property of stretching and rotation as well actually, what you mean by that? So, let us operate this matrix say with respect to a vector P starting from origin let us say, which is coordinate x and y, x 1 and x 2. If I have some point P which is given by a vector from origin to P that is coordinate x 1 and x 2, I multiply that with A, then I will get y 1 and y 2 for which I will get a point Q let us say y 1 and y 2 can be very different from x 1 and x 2.

So, in that sense this vector what you see here and this vector what you see there are different in both magnitude as well as direction basically. So, that means this particular matrix serve says both stretching operation because the vector lengths are different, as well as, rotation rotational property, because the directions are also different actually. Now, then we naturally ask a question that, does there exist some combination of some vector with respect to this particular matrix, for which it can only act as a stretching operator? And, there is no rotation actually. If I take this vector x and then multiply with this matrix A, then I will get some other point. But that will be only a stretching operation not a rotation operation actually.

So, those directions whatever direction pops up that is called Eigenvector direction or principle axis and things like that actually. So, those are the direction that are given by eigenvectors essentially and their magnitude are stressing is actually Eigenvalues. And, there is a general, in general this operation this magnitude can be complex no more and thing like that actually, so we will not venture too much into that. But then you say that we are in, we are searching a particular vector x with respect to that particular matrix A that we have, so that this linear operator applied with respect to that vector will give me another vector with only stretching operation, no rotation actually. So, that is the Eigenvector direction and the amount of stretching is nothing but Eigen values.

(Refer Slide Time: 11:20)



Now, relying on this philosophy what do the math says now actually, how do you operate that? So, what are the questions? The question is can we find a direction or a vector along which the matrix will act only as a stretching operator? And, the answer to that is if that happens then what is that it is a stretching operator anyway. So, A X whatever A X is that is nothing but lambda times X, so lambda is just a scalar quantity now. A X is equal to lambda X will give us only stretching operator actually, that is what we are talking here, A X is equal to lambda X and if that happens then I can take this one to the right hand side and then we will tell this is lambda I minus A into X. Remember, lambda is a scalar but we are talking about a vector and matrix operation here, so we take lambda times I actually.

So, lambda I minus A equal to zero times X equal to 0. So, if that happens, then if this equation is true for a non trivial solution for X, that means this equation is anyway true for X equal to 0. But we are not interested in that, we are not interested in origin point actually, we want a direction actually anyway. So, for a non trivial solution this will have a solution provided this matrix the coefficient matrix is singular. And, you may have infinite solutions by the way. If the matrix the coefficient go singular then we can have actually infinite solution because like on the rank of the matrix reduces in that sense actually.

But anyway coming back to that, this non trivial solution requires that we have this determinant equal to 0. And, that is how we typically see in under graduate texts and all that we deal with this equation common because characteristic equation anyway. So, given a problem we directly can jump into this equation, lambda I minus a determinant equal to 0 and try to formulate a polynomial in lambda. If A is n by n, remember all these are nicely defined for square matrices actually, therefore non square matrices Eigen values, eigenvectors are not defined. So, lambda I minus A is also a square matrix and you take a determinant if A is n by n, then this polynomial will be n th order polynomial.

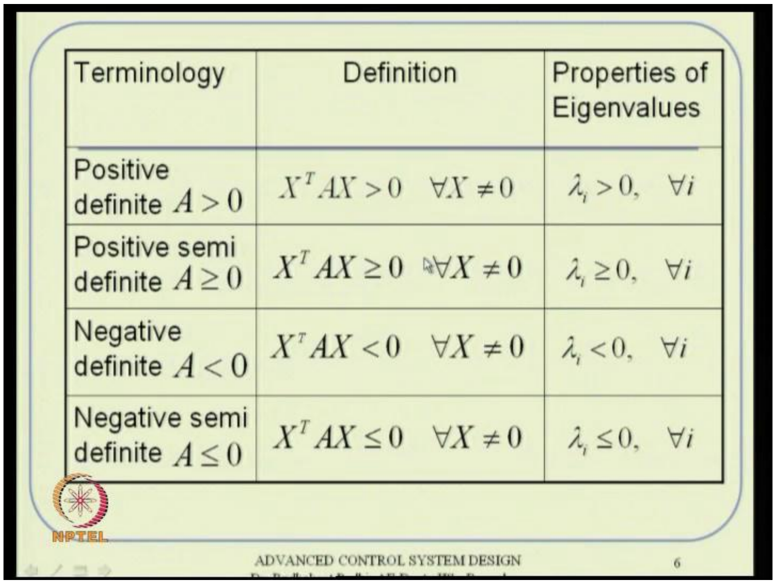
So, we will give something like a n values for lambda, they may be repeated, they may not be repeated, they can be partly repeated, partly non repeated, they can be partly complex conjugate as well. Assuming, that A contains only real values, then this characteristic equation can partly take complex conjugate solutions as well actually. And, then whatever corresponding X that pops up that is common, that is known as eigenvectors. So, the solution each of, for each of the solutions actually of lambda, we will get a corresponding X and that will give us the corresponding Eigenvalue Eigenvector pair actually.

So, in other words going back to that diagram that need not be only one direction, there can be multiple directions actually, for which this equation is satisfied. And, then that will be given by this equation and you will have n sources of pairs Eigenvalue Eigenvector actually coming out of these two equations. And, also remember that eigenvectors do not have a specific magnitude for say, because the moment this is the determinant is 0 we have infinite solutions for X. So, any value, any particular the vector is an Eigenvector by definition, we really do not worry about the magnitude of Eigenvector. Eigenvector is typically a direction, but for our further discussion and all normally we normalized Eigenvector, that means by any Eigenvector we can take a normalized quantity and tell that is my eigenvectors actually. So, that is typically we do that way, but by definition Eigenvectors do not have a magnitude actually.

So, where are the utility? Utility is heavily there in stability and in control analysis and synthesis that is why you are studying this. This is also there for model reduction and it will give you principal components and all the insignificant components you can neglect, that will be coming out from the magnitude of Eigenvalues actually. Then, in case for example, if it is a symmetric matrix we will see it later, that the Eigen values are guaranteed to be real and then you can arrange them in a descending order and thing like that. Then, you can truncate that spectrum and tell up to that I will take and beyond that I will neglect, that kind of concept is called model reduction actually.

This is also used for principal component analysis, that means if your data is kind of scatter around and you want to find out the principal directions around which the data is scattered. And, you want to neglect the data direction thing like that, there also you want to have something called principal component analysis and there also Eigenvalue Eigenvector plays a heavy roll. So, they are everywhere, they are there in stability control, they are there in model fraction, principal component analysis, stress vector analysis, a structural analysis, their applications in fluid dynamics everywhere actually. So, but we are primarily interested because of stability and control property, it is embedded in that actually.

(Refer Slide Time: 16:12)



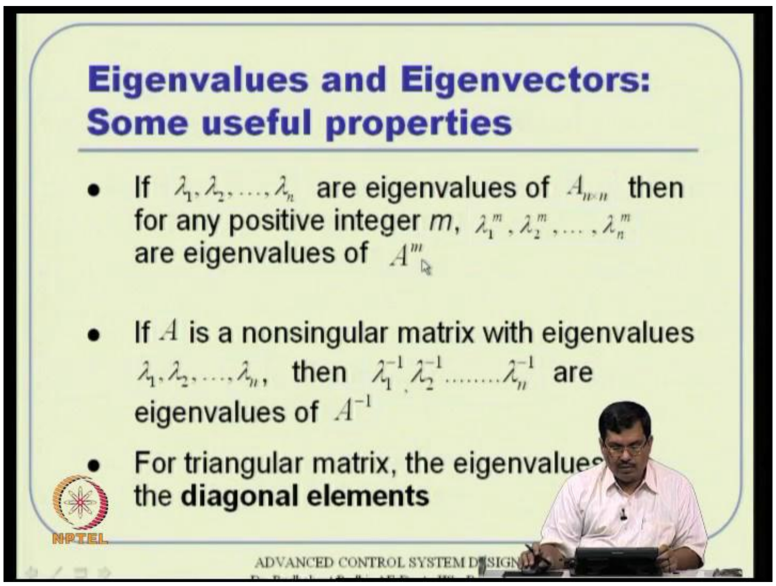
Terminology	Definition	Properties of Eigenvalues
Positive definite $A > 0$	$X^T AX > 0 \quad \forall X \neq 0$	$\lambda_i > 0, \quad \forall i$
Positive semi definite $A \geq 0$	$X^T AX \geq 0 \quad \forall X \neq 0$	$\lambda_i \geq 0, \quad \forall i$
Negative definite $A < 0$	$X^T AX < 0 \quad \forall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

Next terminology next concept what are called as positive definite negative, definite matrix, like that actually. So, a matrix is called positive definite provided this property holds good actually. That means if you take any X, any vector which is a non 0 vector and I do this algebra that is A transpose AX, that is always guaranteed to be strictly positive. And, this is obviously a scalar quantity remember that X transpose A X and these are all for square matrices actually. So, if I take any square matrix and multiply with appropriate dimension vector X to the right and X transpose to the left, for any X I take. I do not care about that magnitude as long as that is non 0, I am guaranteed to get a positive number. That is called positive definite matrix if that happens actually.

And, similarly, if it is happens to be like greater than equal to 0, then it is called positive semi definite matrix and similarly, if it is strictly less than 0 it is negative definite and less than equal to 0 it is negative semi definite. And, notation wise you can still write it, but this is simply a notation you know that. Because A is a matrix and A greater than 0 does not mean element each of the elements of A greater than 0, that is not the meaning actually, the meaning is in this sense. So, many of the analysis in control system we can see that A is greater than 0, that means simply what it means is A is a positive definite matrix actually, in the meaning is like that. And, it is a natural consequence that it is actually very tightly related to the Eigenvalues and if it is a positive definite matrix it turns out that all the Eigen values of this matrix are guaranteed to be positive numbers.

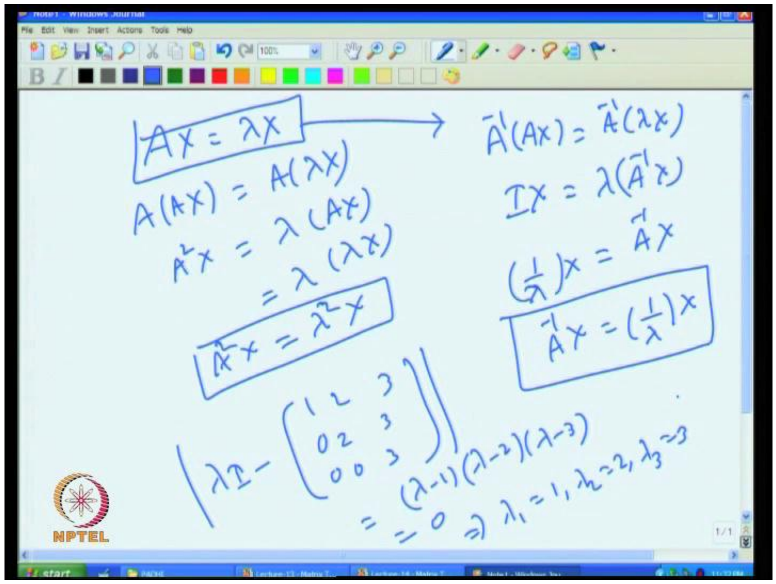
So, if you have studied this definition from Eigenvalue property point of view, my suggestion is these are not definition, these are all consequences actually. But sometimes they are also taken as definition because these are if and only if results actually, this is not one sided results. That if this happens that is going to happen and if this happens that is also going to happen. So, sometimes these properties of Eigenvalues and definition they are kind of interchanged, but in strictly speaking these are definition and these are properties actually of those matrices. You just remember that.

(Refer Slide Time: 18:35)



So, now next we will study some of these useful properties of Eigenvalues and eigenvectors. So, let us study one by one actually. So, what are we telling here? The first thing is if lambda 1 to lambda n are Eigenvalues of a matrix n by n, then for any positive integer m this powers of these Eigenvalues are also Eigenvalues of this A to the power m. It is a very beautiful property in that sense because if I know the Eigen values of A matrix, I really do not have to compute for A square, A cube, A 3 A forth all sort of things actually. I can simply take their powers and I will directly get the Eigen values of that. it is rather easy to easy to solve also. Let us show that in a very quick way.

(Refer Slide Time: 19:27)



So, you have A lambda as a Eigen value, so you have your equation is satisfied this way, this is by definition anyway. So, I want to show that lambda square is an Eigenvalue for A square and then you can generalize that for A cube, A forth and all that actually. So, for that what I will do? I will just pre multiply both sides by A, so I get A times lambda X, remember vector matrix is this pre multiplication and post multiplication has to be done carefully. So, then this one is A square into X. However, lambda is a scalar quantity, so I can take it outside that and then lambda into A X that I will get. And, then this is A X is equal to lambda X, so I will take this is nothing but lambda X and hence it is lambda square X. So, what I getting here? A square X is nothing but lambda square X.

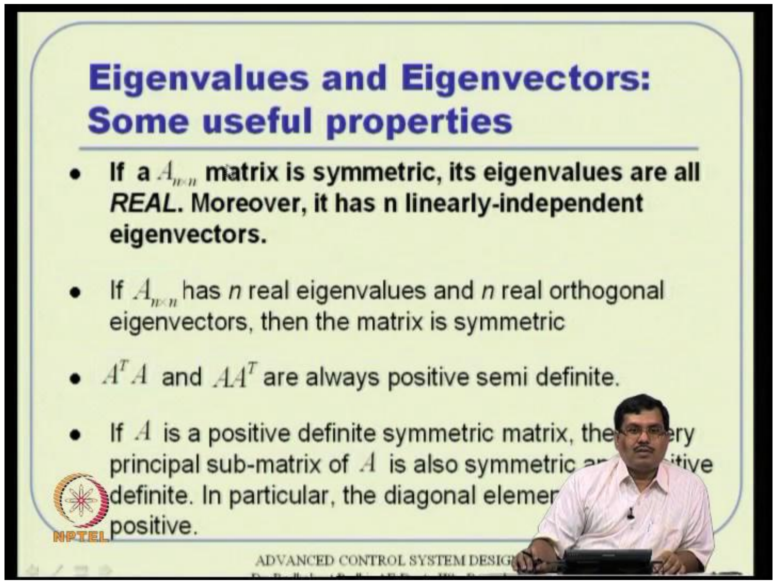
So, obviously by definition again what we get it here, this tells me that lambda square is nothing but an Eigenvalue for A square. That tells me very clearly actually about that, so what you get here. So, this A X equal to lambda X, so the similar expression is there. Now, if you multiply one more time here, this equation if we multiply one more time, you get A cube X is still lambda q times x. so, like that you can continue and then tell for any integer power I will get lambda to the power m is actually as Eigen value for A to the power m. And, beautiful thing is as a consequence, as a as a corollary eigenvectors remain same. You see that same x appears here, same x appears here. So, Eigenvalues will take their powers and Eigenvectors will remain same actually, so that is a nice property.

Now, the similar thing is also true for minus one as well and remember minus one means inverse actually. A to the power minus one that means A inverse, so if I know the Eigenvalues of a matrix then all that I have to do is 1 by lambda 1, 1 by lambda 2 all that I can take and then claim that these are actually Eigenvalues for a inverse as well actually. So, how do I do that? That is also simple starting from this equation, I will start from this equation until I, this time I will multiply with a inverse rather. So, I will get A X equal to I mean A inverse into A X is equal to A inverse lambda x. So, if I plot, if I do the simplification A inverse into A is as identity, so I times X is equal to lambda times A inverse of X and identity X is I times X is A X obviously. So, what I get? I will get 1 by lambda times X is equal to 1 by lambda times X is nothing but A inverse X.

So, the same equation if you just rewrite the left side into right side, right side into left side, what you are getting? A inverse X, A actually 1 by lambda times X. So, this equation tells us that 1 by lambda is an Eigenvalue for A inverse actually and again as a corollary Eigenvectors remain same. So, they are rather easy, I mean they are not difficult as even though the results are very powerful, these results are not very difficult to solve at all actually. Now, the third one for a triangular matrix the Eigenvalues are diagonal elements, these are very clear from the definition itself actually. If you go back to the definition and tell this is what I have, suppose if I have a triangular matrix then this equation is nothing but lambda minus the diagonal element 1, lambda minus the diagonal element 2 like that actually.

So, you can if you want to just have a example. So, if I take let us say 1 2 3 and then let us say 2, I mean 0 2 3 and 0 0 3 let us say, something like this matrix. Then, the characteristic equation is lambda I minus A determinant actually. And, that polynomial will give me lambda minus 1 into lambda minus 2 into lambda minus 3 and if I equal it to, if I equated to 0 then all that I get lambda 1 is 1 lambda 2 is 2 and lambda 3 is 3. So, the characteristic equation only diagonal elements will pop up actually, other things will not, will not appear. So, that is why it is more powerful actually. Anyways, that is the result why it is, why tell that for a triangular matrix Eigenvalues are diagonal elements and there is a, I mean it is very clear that diagonal matrix is also a triangular matrix. And, hence for diagonal matrix the Eigenvalues are nothing but diagonal elements actually.

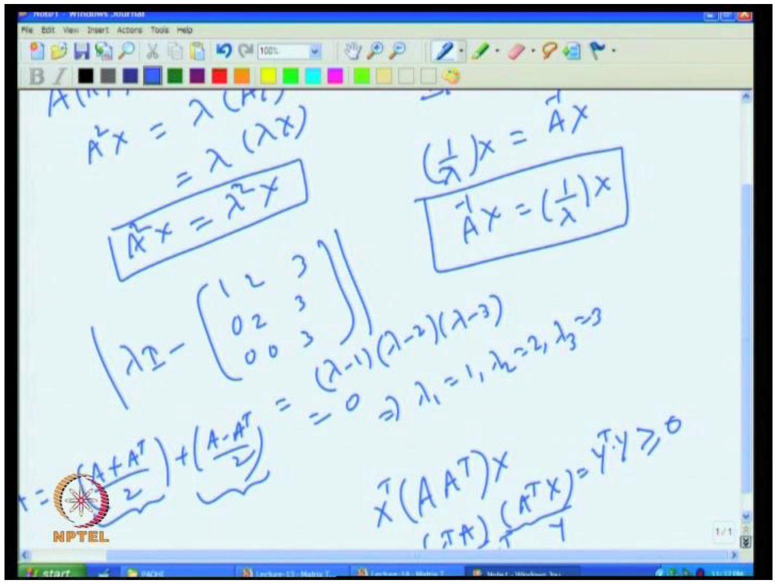
(Refer Slide Time: 24:04)



Now, the next big property which is very useful to us is something like this which tells us that if a square matrix is symmetric, then the Eigenvalues are all going to be real, it is a very very powerful result actually. All if I know that a matrix is symmetric, then I am for, I mean for sure I can say that Eigenvalues are guaranteed to be real actually. So, the complex conjugate another thing will not appear, the Eigenvalues are not going to be complex at all actually. Then, on top of it moreover it will have n linearly-independent eigenvectors as well, so it is a very very powerful result. And, hence most of the time given a choice we would like to formulate a symmetric matrix. If you have to ever formulate a matrix then typically we will formulate a symmetric matrix in our problem formulation actually.

So, suppose a matrix is given to us, we do not have a choice, we have to work with whatever we have. But if it is even in that case, we can actually talk about decomposing into the something like, suppose A is a non square matrix.

(Refer Slide Time: 25:16)



Then, we can decompose that something like A plus A transpose by 2 plus A minus A transpose by 2 as well. This is supposed to be symmetric, this is supposed to be kind of q symmetric, what is called actually A. So, this is still we want to do because if this component is negligible, that means if this q symmetric component part of it is not that significant. Then, we will still left out with an approximate matrix which is still symmetric matrix actually. So, that is a good property that we will have to explore it actually. And, some of these proofs and all, you can if your time permits we will see it in a class otherwise you can read it from standard text books actually, that is also possible to do.

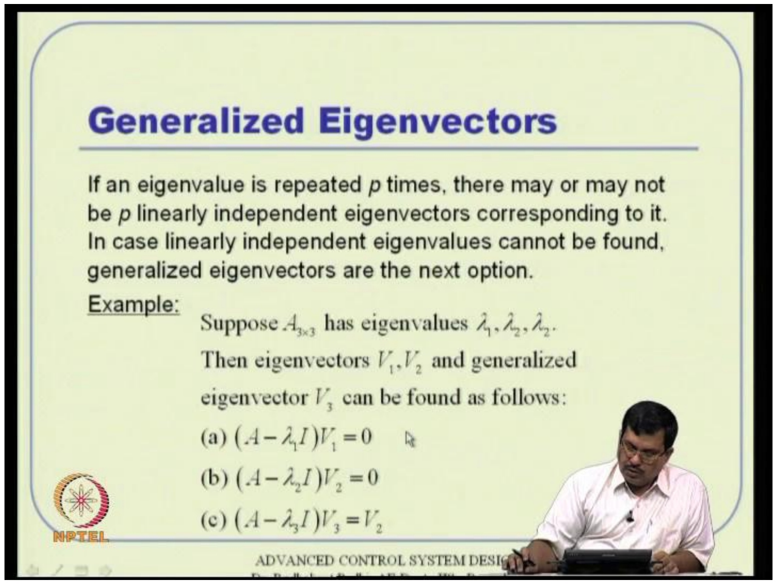
And, the next property, next big property is if a square matrix has n real Eigenvalues and n orthogonal Eigenvectors, then the matrix is guaranteed to be symmetric as well. That means this property is actually a converse of this property. So, whatever you see there, if a matrix is symmetric then the Eigenvalues are all going to be real and it has n linear-independent eigenvectors. Now, the converse property tells us that if a matrix is a n Eigenvalues and n real orthogonal Eigenvectors, remember these are not necessarily linearly independent eigenvectors only, they have to be orthogonal as well. If it is n orthogonal Eigenvectors then the matrix is guaranteed to be symmetric, that is the converse theorem actually.

Now, the next property is what is very handy many times is, if you have a matrix something like A transpose A or A A transpose this is guaranteed to be positive semi definite always and it is not that difficult to solve it either actually. So, how do we show that? We will see, if you have a matrix something like A A transpose let me do that or A transpose A either way, then what you do? You multiply that with let us say X transpose and X actually, that is what you really want.

So, and then you can tell this is nothing but X transpose A into A transpose X and one is the transpose of the others. So, this is suppose this is Y then this is supposed to be Y transpose, so and then this is nothing but this is equal to Y transpose Y that means it is something like Y 1 square plus Y 2 square plus Y 3 square like that actually. This particular quantity what you see Y transpose Y is Y 1 square plus Y 2 square plus Y 3 square all that and that is guaranteed to be greater than equal to 0 actually. This particular thing is guaranteed to be greater than equal to 0 because of that property. So, that is why this is positive semi definite actually. And, you can show that for the other one also, if it A transpose A that is also easy to show that actually.

Now, next property if A is a positive definite symmetric matrix, then every principal sub-matrix of A is also symmetric and positive definite. In particular the diagonal elements of A are always guaranteed to be positive actually. So, this is also remember that, if it is a positive semi definite, I mean positive definite matrix then the diagonal elements are guaranteed to be positive because of this property. So, if you see a matrix where diagonal elements one of the diagonal matrix is negative, then you can right away tell that the matrix is not going to be positive definite. So, that is consequence of that.

(Refer Slide Time: 28:57)



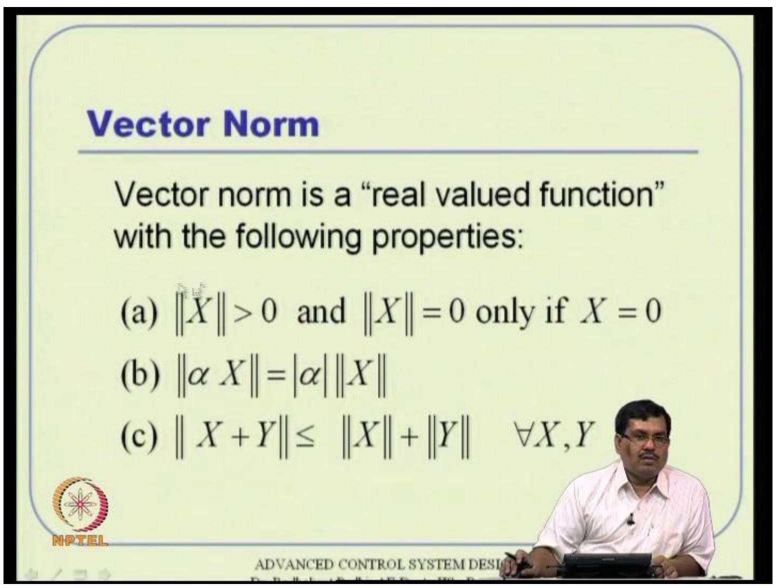
Now, we will go to the next property and this is something some concept call generalized Eigenvector actually. So, as you remember that if you have independent Eigenvalues like linear-independent Eigenvalues, then you are guaranteed to get independent Eigenvectors as well, so non-repeated Eigenvalues what I mean. If you have non repeated Eigenvalues, then you are guaranteed to get linear-independent Eigenvectors because of that property that I mean. Now, if you come back to that then ask questions, what if the some of the Eigenvalues are repeated actually? Do we still guarantee that?

Now, it turns out that if some of the Eigenvalues are repeated, then you may or may not get independent Eigenvectors actually. And, if you get independent Eigenvectors that is well and good and if you do not get independent Eigenvectors then we would like to have some vector which is actually rather close to Eigenvectors. It may not be exactly Eigenvector, but the properties may be very close to the Eigenvector actually. And, some precise definitions are also available for general Eigenvectors. But in a way we want some Eigenvector which is very close to the Eigenvector in some way actually. How do you compute that? If you have let us say you have lambda 1 and lambda 2 lambda 2 that mean lambda 2 is repeated twice.

Then, for lambda 1 and lambda 2 you will get v 1 and v 2 anyway, so that is fine. But for the third lambda 2 which is also same as the second lambda 2, you suppose there is no n linear-independent eigenvectors then you formulate this sort of an equation where the right hand side is not equal to 0, but it is equal to v 2 rather. So, then whatever solution you will get v 3 that will be a like a generalized eigenvectors say actually, that is not Eigenvector truly speaking. But it is closed to Eigenvector in some so that, if say that in a itself on delta sense there is a limiting sense behavior and all that, that is available in Kailath book actually. If you want to see that Thomas Kailath book, so you can refer to that and tell that is the property that will satisfy in a limiting sense, it will become an Eigenvector actually.

So, the equation that is needed to satisfy is like this actually and also suppose this lambda 2 is repeated one more time, then the next equation will be this into v 4 is equal to v 3. So, this is v 3 is equal to v 2, the next one will be v 4 equal to v 3 here, that way you will compute v 1 v 2 v 3 v 4 part of them will be really Eigenvectors, real Eigenvectors and part of them will be generalized Eigenvectors. And, these Eigenvalues, Eigenvectors set are actually very useful in matrix transformations especially similarity transformation. We will see that in the next class actually.

(Refer Slide Time: 31:45)



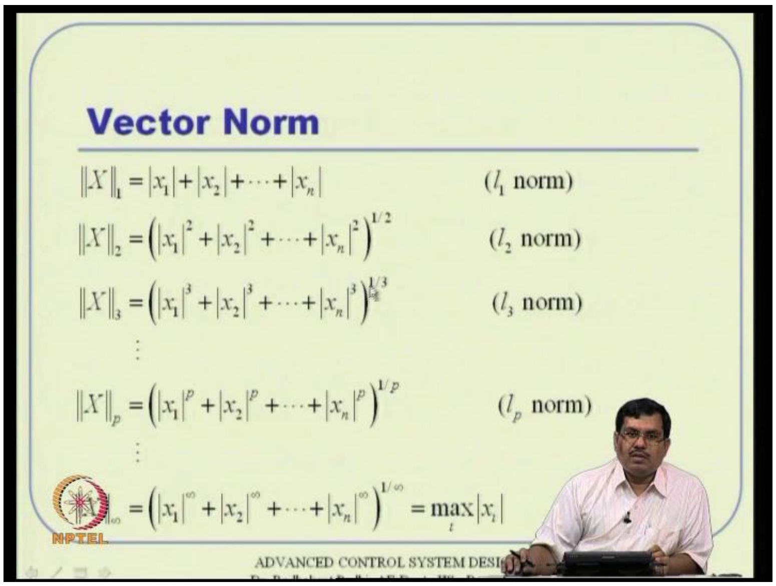
There is a next concept what is useful to us is something called Vector Norm. So, what is a norm? Actually, norm is a concept of distance sort of thing, that means how big is a vector. Suppose, you have a scalar value we can really compare one is smaller than the other, I mean the magnitude wise one is lesser than the other quantity one is bigger than the other quantity like that actually. So, somewhat we also want to know, what I mean the vector sense, how much it really quantify? So, I mean one way to quantify is just Euclidean distance from origin, that what we saw that in one of the slides before that the length part of it actually. That is the Euclidean distance from origin actually.

So, if you really take that distance that is nothing but square root of x 1 square plus x 2 square, that happens to be a second norm actually. And, somebody can always argue that hold on a second I will not do it that way, let me let me travel from origin to that point using this way, I will go this way first and then go that way. I will still cover some distance actually. So, this is parallel to x axis I will go and then parallel to y axis I will travel. So, if I just take the summation of this distance plus that distance that happens to be what is called as l 1 norm actually. That is also distance quantity by the way, that is also going to be a positive quantity, that also signifies how we give a vector and that happens to be a norm quantity as well actually. So, like that we can define various ways and the formal definition turns out to be something like this.

Norm vector norm is actually, remember, it is a vector norm that means it is defined for a vector not for a matrix in general. So, X is a vector and the norm of that vector is a real valued function, that means actually it is the ultimate number the norm is essentially a scalar number. That real valued function will satisfy these three properties actually, what is that? Norm of a vector is guaranteed to be a positive quantity; it is never going to be a negative quantity as long as X is non 0. And, if X is 0 then the norm is guaranteed to be 0; that is the first property. So, if as long as it is a non 0 quantity, non 0 vector that is you are moving away from origin in any direction whatever direction is that, then the norm of that quantity is guaranteed to be a positive number and it is very comfortable to the distance travel concept. So, no matter what direction you travel the distance travelled is guaranteed to be a positive number anyway. So, that is that is kind of a similar thing there.

Now, if you do a scalar multiplication of that vector and then take a norm of that, then it should satisfy the absolute value of that number into norm of that vector that is to satisfy that actually. Then, quantity c that the property third property is something called triangle in equality and that tells me that if I add two vectors of similar dimension of course, and then I take a norm that will satisfy a less than equal to property that way. No matter whatever is my X and Y, if I take them X plus Y and then take a norm of that, then it is guaranteed to satisfy this less than equal to norm of X plus norm of Y. So, any quantity, any function that you define for norm has to satisfy this, these three properties actually.

(Refer Slide Time: 35:14)



And, if and obviously satisfying these properties we can define these quantities actually. That if I simply define the absolute of x 1 plus absolute of x 2 plus absolute value all the way up to x n, then it will satisfy all these three properties anyway. Absolute values, all the absolute values are positive quantities, so it is the summation of that is guaranteed to be positive quantity also. It is going to be 0 provided only if x is 0. So that is, that will satisfy from this definition, it will also get a satisfy from all these definitions actually. And then or you can also verify that the rest of the two quantities are also satisfied, I mean they will also get satisfied with respect to all these definitions actually. And, if you have only this formula then it is something called l 1 norm and this formula is something called l 2 norm, this formula is something called l 3 norm like that actually, you can define any order norm by the way.

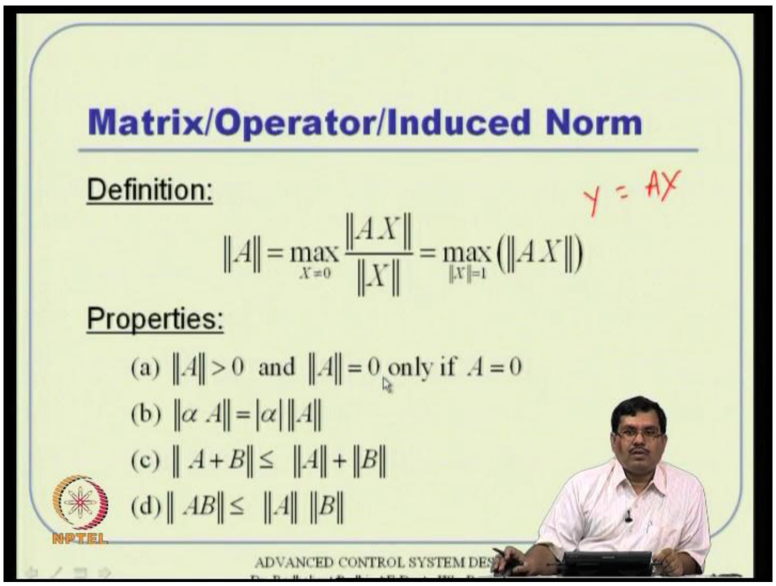
So, p th order norm, l p norm is defined like that so you take absolute value of each of the quantity raise it to the power p, take their summation and I take p th root of all that actually. So, you take first raise the power p then take the sum and take p th root of that actually, positive p th root of that. So, then you will, whatever quantity you will end up with that is called p th norm actually. So, starting from one norm second norm three norm four norm five norm up to infinity norm you can define, by the way. Then, it is so happens that the infinity norm also happens to be just that, if we simply take the maximum of that quantity, whatever quantity you are having x 1 x 2 all that thing if you simply take maximum of that value, that whatever component is the maximum one, then that is nothing but infinity norm actually.

So, if you go back to the diagram and say what is my infinity norm for vector one, then obviously it is going to be absolute value of x 1. And, for vector two it is going to be absolute value of y 2 basically, so like that actually. So, typically in our analysis of control theory, we will require l 1 norm, l 2 norm and infinity norm, we will all work along with this these three lines only. And, there is also a theorem which tells us, little later I will tell probably, that in a finite dimensional vector space all norms are equivalent actually, they are not equal of course. But they are, they share a equivalent property. What does that mean? If you show that second norm is bounded then all other norms will also get bounded actually, that need not happen in infinite dimensional vector space.

But as long as finite dimension vector space, n is a finite number then all norms are kind of equivalent actually, that theorem is a very handy tool which will tell us that is ok, we will not have to worry about each of the terms individually each of the definitions actually. So, most of the time we will either deal with second norm, but you will see that all l q r theory quadratic things and then this Kalman filter all these things. We will talk about h two sort of an idea, sometime we are called h 2 norms also, I mean h constant from I d space and all that actually I will not go too much into that.

And then, either we deal heavily into two norm or sometimes we deal, we like to deal with the maximum quantity of the noise and thing like that, your system has to be stable for maximum noise magnitude that you expect actually, noise frequency and magnitude all that actually. That concept is h infinity concept. So, for h two concept control theory we will require the second norm, but h infinity concept we will require the infinity norm. So, mostly we will do with l 2 norm or l 3 or l infinity norm actually. But remember as long as I show something all the norms are actually equivalent actually.

(Refer Slide Time: 39:03)



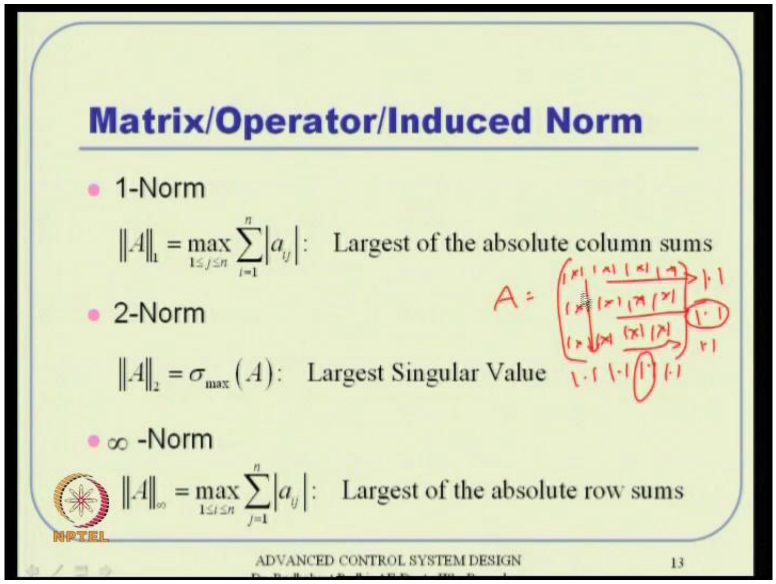
All right, that next it throws to a natural question that we discussed about vector norms here what about matrices actually? That means, if you even if I have a matrix, can I talk about, how big is that matrix? And, actually it is also related to how far I can stretch the vector actually? Remember that. Suppose A X, now remember A X is nothing but something like a, we have seen that in Eigenvalue and Eigenvector, Y equal to A X. And, then A X, well let me Y equal to A X and also we have seen that this is, this also serves something like a stretching operator actually, we are not so much concerned about rotation here, let us say. Now, we will talk about stretching operator, now no matter whatever is my A X, this Y this A times X I want to know how much this vector is getting stretched actually, getting elongated how much it will go elongated actually?

Then, if I say that property I want to extract then let me extract me in that way. I will take any vector that is non 0, I will first formulate an A times X matrix that will give me a vector Y. And, that magnitude of vector Y, I will divide it by magnitude of vector X. So, whatever is that ratio, that magnitude of vector Y divided by magnitude of vector X, will give me a, certainly it will give me a positive quantity, positive number actually. And, that number is nothing but matrix norm actually. And, suppose I have to simplify my algebra, I consider something like I will constrain my norm X to 1. That means I will work around that unit circle or unit ball or unit sphere thing like that on the surface of that, then the all the values of y that will pop up I will take them and then find out what is the maximum value out of that.

So, it is the maximum stretching condition that comes into picture while defining, what is called is matrix norm which is also called as operator norm for obvious reasons. Matrix is a linear operator; it is also called as induced norm actually. So, they are three different norms for different names for the same thing actually. Anyway, so that is the idea. Idea is this matrix norm is nothing but a stretching operator; I want to extract how much it stretches with what vector and the maximum stretching what it can do in whatever direction that is actually my matrix norm. So, remember if I take different X, the stretching can happen in a different magnitude basically. So, I am considering the maximum stretching out of that.

So, what are the properties that it takes to satisfy? Again you see that first three properties are very similar to what we have earlier. But it also needs to satisfy some sort of a triangle inequality in multiplication sense as well. Only addition is not sufficient, it will also satisfy something like a multiplicative inequality, which tells us that A times B norm is less than equal to norm A times norm B. So, that is what it will satisfy.

(Refer Slide Time: 42:13)

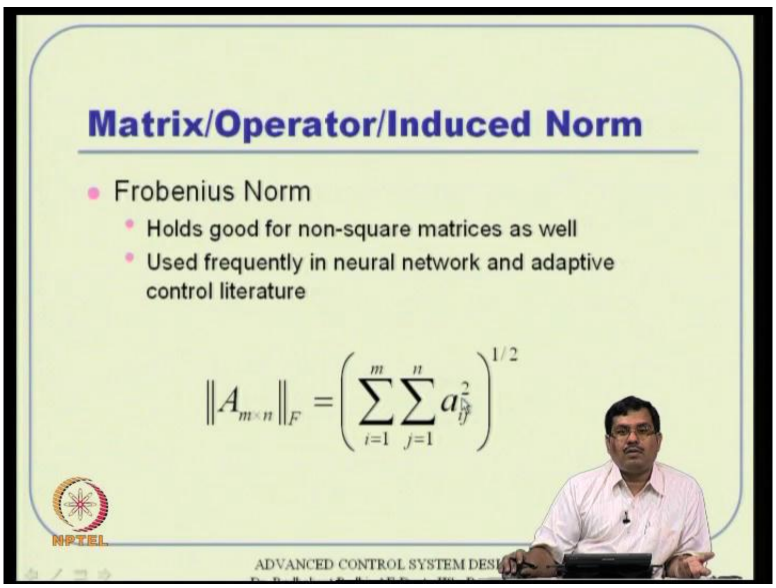


And, then definition wise something define like that, it tells how what is this first 1 norm 2 norm and infinity norm. Again, as I told these are the three quantities that we will worry most of the time. And, let us talk about 2 norm which is very popular which we talks which is nothing but maximum singular value. And, singular value is something that I will take you through in definition sense and all that. So, this if you take all singular values and find out the maximum singular value that is nothing but the second norm actually. So, that is kind of a case fairly straightforward in that sense, as long as you find the singular values basically.

Now, what is 1 norm and second norm, these are defined something like that, mathematical notation wise they are like this. However, English words sense these are nothing but largest of the absolute column sums. And, this is nothing but largest of the absolute row sums actually, so what do you mean by that? So, let us go to that and tell there is a matrix out here, a matrix there is all numbers actually spread out everywhere. Now, what do I do? I take absolute values of all the terms actually, whatever values are there. The if I once I take absolute values of that, then I can sum it up across rows, all of them sum as then I will get a number here, number here, number here. And, I will consider the biggest one here; let us say this happens to be larger than first and third. Then, that happens to be infinite norm because it is nothing but largest by definition, it is nothing but largest absolute row sum actually.

Now, if I do the similar operation column wise, if I can take this one and then find a number, this one and again this column wise find a number and find a number and tell this is the maximum out of all that. And, then that happens to be my largest absolute column sum and that happens to be my one norm actually. So, the one norm is nothing but largest absolute column sum and the infinity norm happens to be largest absolute row sum and it is rather easy to remember, I do not have to get confused unnecessarily. You can remember letter 1 is vertical and letter infinity is horizontal, the way you write letter infinity is horizontal and the way you write letter 1 is vertical. So, that is what it is very clear to see that way.

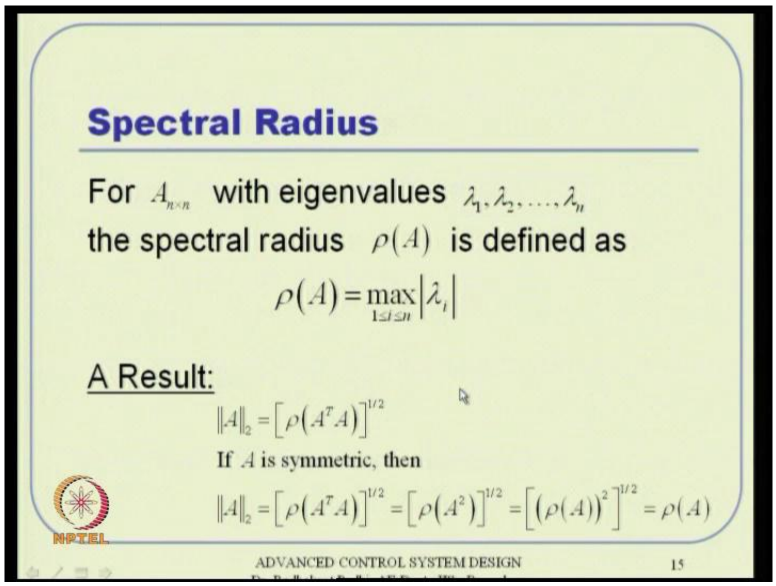
(Refer Slide Time: 44:43)



Now, going to the induced norm I mean there is also another concept called Frobenius norm, which is also useful in note of this matrix theory analysis, neural network analysis and all that which is nothing but almost like a second norm for vectors actually. What you do vectors, all these you take square square square, add it up everywhere and then just take summation of everything and then take square root of that first square root of that. And, similar concept if you do, you just blindly just square of all the terms in the matrix and after taking all the squares you just add them together. And, then take the final quantity you take square root of that. That will give you Frobenius norm actually.

It will also satisfy this basic property that you talk about and hence it is also a norm. So, that particular thing is also used for neural network training and all, we will see some of this adoptive control, because none of the ways in the matrix you really want to go unbounded actually. This is this will satisfy certain nice properties actually.

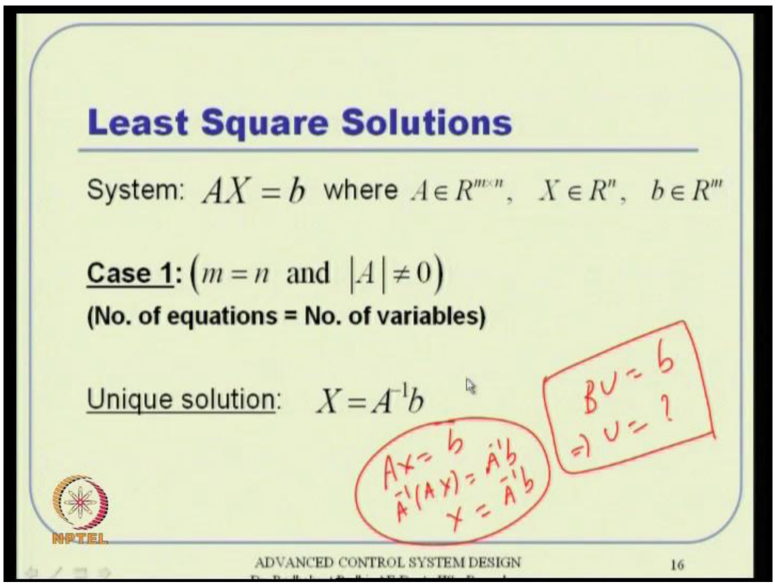
(Refer Slide Time: 45:36)



Now, the next concept which calls as spectral radius. Spectral radius is nothing but something like, maximum of these absolute values of lambdas actually. So, if I take Eigenvalues of a matrix n by n and then simply take the absolute values of that Eigenvalues, remember if even if there is a complex Eigen value I can still talk about a magnitude of that quantity, that complex number. And, then I will find out what is the maximum absolute value of that Eigen value. That is nothing but the spectral radius of that matrix. So, Eigenvalues remember they give the quantity the magnitude of stretching basically. So, I want to know what is the maximum stretching that can happen also in a way. So, that is nothing but spectral radius actually.

So, now a standard result tells us that spectral, that is the 2 norms. 2 norm is related to the spectral radius that way and if A is symmetric is, if you talk about symmetric matrix then A 2 is nothing but that by definition. And, if it is symmetric you know A A transpose is also A that means is A square. And, then it will happen that you this square and remember this after all is an Eigenvalue actually, the property is related to Eigenvalue. An Eigenvalue of A square is Eigenvalue of A to the power 2; we have seen that one of the properties. So, you can use that way and then come across, this is what is going to be my spectral radius actually. So, the second norm is nothing but the spectral radius of the matrix also, that is very easy to show that.

(Refer Slide Time: 47:18)

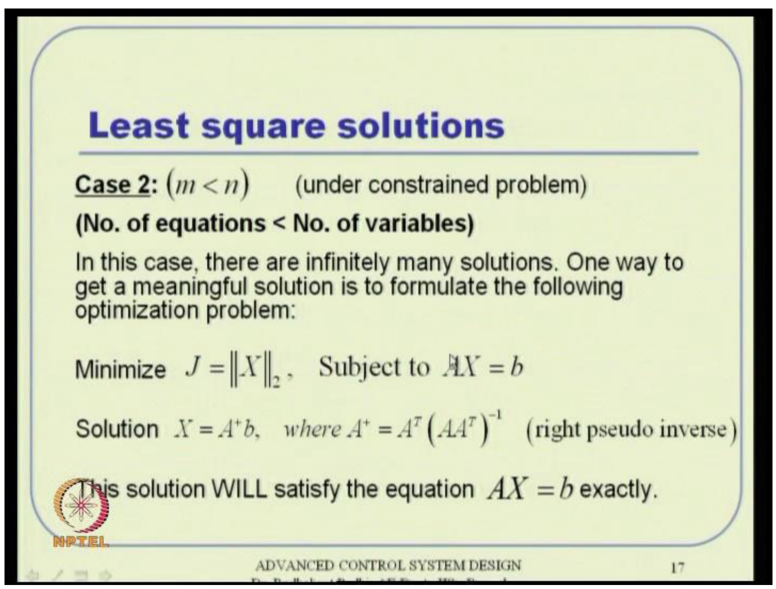


The last concept in this lecture probably we will talk about what is called as least square solution. And, this least square solution is a concept which is very close to the, like solving a system of equation actually. So, we are concentrating, we very repeatedly we will come across this equation. Suppose you I mean in a (( )), in a control theory we will most like the time we will land up is something like A matrix some b matrix something as nothing do with the control influence matrix, but some matrix times U is equal to some function actually like some vector V.

Now, how do you solve for control actually? That is the question that we will typically land up with many many times in control design actually. So, we are generalizing that and tell if this is the equation that is a primary importance to all and then see what all possibilities can happen actually. And, remember X is a n dimensional vector and b is a m dimensional vector in general. That means the number of equations are actually m and the number of variables that we are talking about is n actually. Then they may or may not be same. So, obviously it will arise three cases, now we will study case by case. Now, case 1 is something like, we will concentrate m equal to n, then what happens? That means, number of variables are equal to the number of equations and on top of that we are assuming the determinant of A is not equal to 0.

Then, the solution is very straightforward, we have seen that in I mean many many times. That it is nothing but X equal to A inverse b, so if you multiply pre multiply both sides by A inverse because start with A, A X equal to b. So, pre multiply both sides by A inverse A X is equal to A inverse b. And, then you tell this is nothing but identity so X X equal to A inverse b. So, that is very very straightforward actually, so that is nothing there actually. As A inverse is always possible because determinant of A is not equal to 0, A is a non singular matrix, so that is a unique solution and you have seen that many times actually already.

(Refer Slide Time: 49:29)



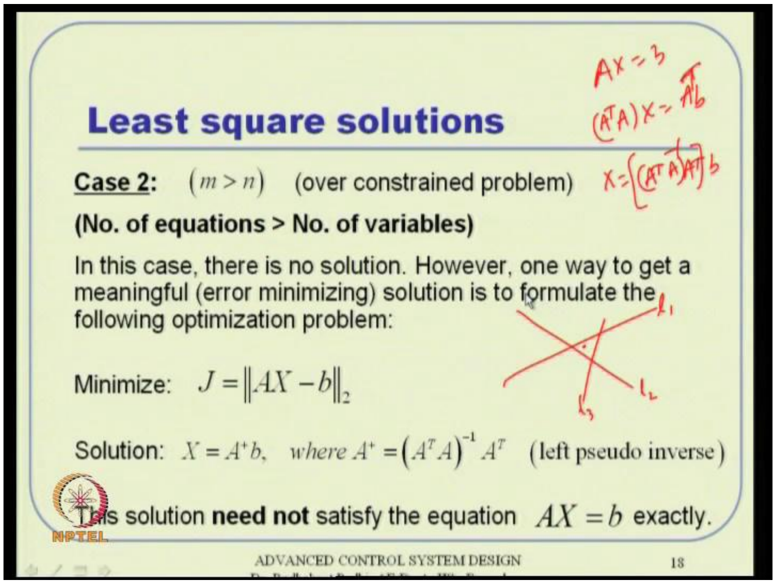
Now, the interesting case turns out if that matrix is really non square, that means m is either greater than n or m is less than n. so, let us study first case, case 2 where m is less than n, that means it is actually under constrained problem. Remember, n is more, n is more means number of variables are more and m is less means number of equations are less actually. That means really it is in under constrained problem, you have more free variables to satisfy this equation. So, obviously you will have infinite solutions actually. So, out of these infinitely many solutions you can try to find a solution in various ways, you can I mean sometimes we force some of the components to be 0 and then , we will enforce that those components to 0 and find out a solution for rest of the thing that is possible.

One other way of telling that let me find out a norm minimizing solution actually. That means I want a solution which will satisfy this equation exactly, but on top of that it will also minimize this second norm of this vector actually. That means if I take the distance from origin, the distance is going to be minimum actually. Yet, I will satisfy this equation exactly, remember that there is no approximate satisfaction; this equation will always be satisfied actually. So, the solution happens to be, I mean if you really carry out this optimization and thing like that, the solution happens to be nothing but A pseudo inverse b, where the pseudo inverse is defined right it is called right pseudo inverse by the way, it is defined something like that.

And, then the details and all you can see in the matrix theory by Ortega, I mean that is a very nice book actually, you can see some of these things there actually. Many of these concepts I have taken from there from that book actually. And, then this solution happens to be a right pseudo inverse if you apply with that this formula, then it will it is nothing but that solution actually. And, remember this solution will exactly satisfy this equation. That means we are not compromising the solution nature actually, solution is guaranteed, on top of that I am getting a little more actually.

And, this concept will be useful to us later when we talk about optimal dynamic inversion. That is where we will also formulate a norm minimizing solution for the control variable; you will have a more control variable then necessary. Then, we will have formulated as problem where we want a solution which will satisfy the equation exactly, on top of that we will have control minimizing solution actually. So, those things will be needed later actually. So, this is the concept where you can see that I can get a solution which is also minimize this, which can also minimize this cross function actually. So, that is my, that is the solution that I am talking.

(Refer Slide Time: 52:13)



Now, what about the next case? That means, if the other case happens that means the number of equations are more than the number of free variables. Obviously, we are talking about a over constrained problem, the free variables are less the number of equations are more actually. So, unless some equations are a linear combination of the other equation, then in that case you can eliminate that equation. But if you assume that none of the equations are having that property that means all the equations are justifiable linearly-independent decoupled equation like that actually. Then, you are caught with a situation where you cannot find an exact solution, where all the equations you cannot satisfy all the time actually.

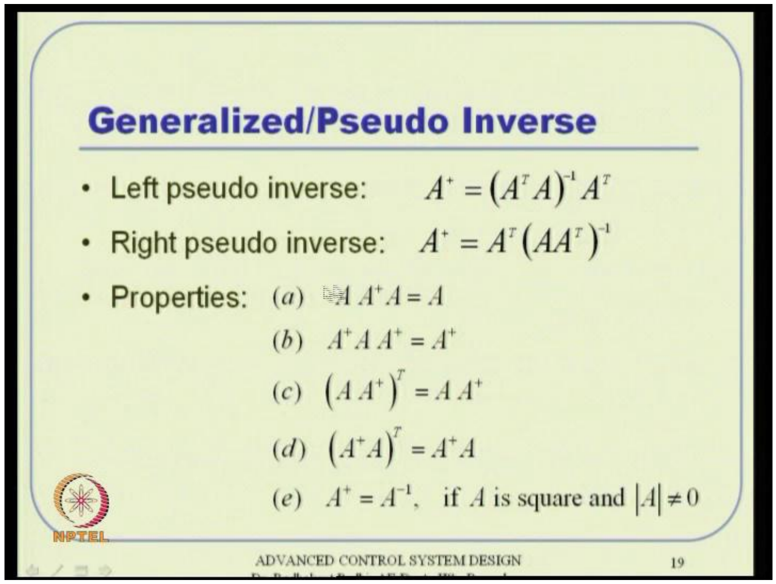
So, very natural way of asking that the next question is can I find some solution which will satisfy most of the equations in an approximate way, all of the equations in an approximate way actually? That means, if you talk about a pictorial idea sort of thing suppose you have two lines then obviously the solution is the point actually. Now, if you have three lines they need, the third line need not pass through that, it can pass through something like that. Then, you are trying to find out a solution which will be close to all the three points actually, whatever three points we talk about right each of the solutions. Suppose, this is let us say this is l 1, this is l 2 and this is l 3, so, obviously l 1 and l 2 have a solution, this is l 1 and l 2 solution, this is l 2 and l 3 solution and then this is l 1 and l 3 solution.

We are not interested in any of that, rather what you are telling is we will find some solution which will very close; I mean which is rather close to all of them actually. So, then that sense we will tell, we will let us go ahead and minimize this particular quantity A X minus b second norm of that we will try to minimize that quantity directly. Remember, this cross function is different from this optimization problem, here we want to minimize the norm directly subject to this constrained. But here you just want to minimize this one as A 3 optimization problem, there is no constrained actually.

And, interestingly this also happens to be a pseudo inverse solution, where the pseudo inverse this time is defined like this. And, this is called as left pseudo inverse actually; it is not right pseudo inverse anywhere. And, so the solution formula if you see this is right pseudo inverse this is left pseudo inverse actually. Now, this as I told the solution need not satisfy the equation A X equal to b exactly, however it will satisfy all of them in an approximate sense actually. And, it is also I mean it is interesting to see that this equation can be derived from a different perspective also.

Suppose, you have A X equal to b then you simply multiply both sides with A transpose let us say, A transpose A X equal to A transpose b and then you take A transpose A is non singular let us say. Then, you actually X equal to A transpose A inverse A transpose b, this is nothing but the same matrix that you see here actually. So, that is also possible to do that. And, there are dimensions if you saw a vector space analysis and thing like that, you are actually working in a radius dimension actually, you are not working in a original dimension. But you are working in a sub space not the full space actually, those connotations those, I mean implications are there actually.

(Refer Slide Time: 55:27)



Now, what is the pseudo inverse or a generalized inverse? It is either it is a we talking about a left pseudo inverse which is defined that way or you talk about a right pseudo inverse which is defined that way. And, all this both of the pseudo inverse will satisfy all of these properties. The very nice thing is, if A happens to be a square matrix and A is nothing but like A is non singular, then this pseudo inverse is actually an inverse. No matter what, whether you take left or right if A is a square matrix and determinant of A is not 0, then pseudo inverse is actually inverse. That is why the name pseudo inverse comes actually.

And, then you talk about all other property what you see, will almost feel as if it is an inverse. Suppose, if you just take an inverse let say, so A inverse is identity, so I left out with it. And, similarly if it is a A inverse is a identity we are left out with A inverse. So, it will feel almost we are working with an inverse, so that is why it is not really an inverse, this a different formula it is not inverse of A. But it is very closely to an inverse actually. So, these are called generalized inverse or the pseudo inverse actually, that is the property.

Besides, it is many usage I mean we will see one of the usage as I told we will see that again in the optimal dynamic inverse, when I talk about that. And, there are other other of the many many things, what you talk this property is very useful in optimization problems actually. Wherever you see optimization problem and if it is over constrained optimization or under constrained optimization that is what typically it happens, then it happens to be like that is the requirement that it pops up actually. So, with that I will stop this class. Thanks a lot.