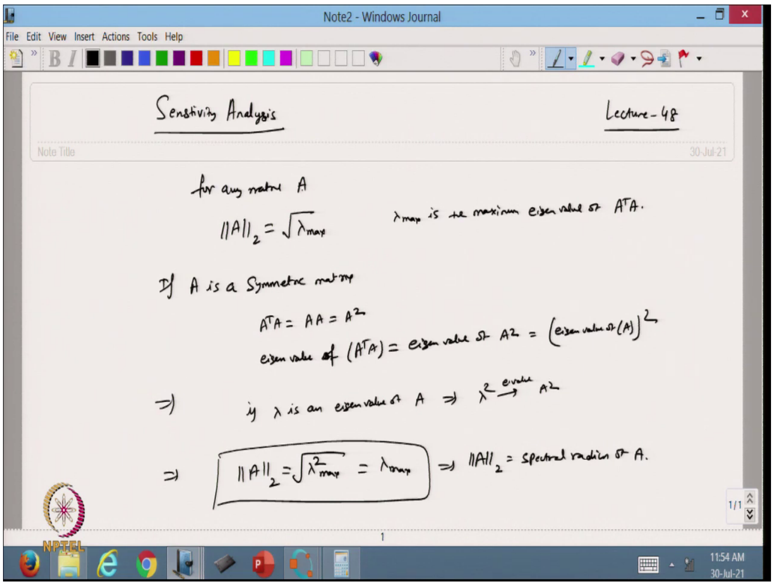
Matrix Computation and its applications
Dr. Vivek Aggarwal
Prof. Mani Mehra
Department of Mathematics
Indian Institute of Technology, Delhi

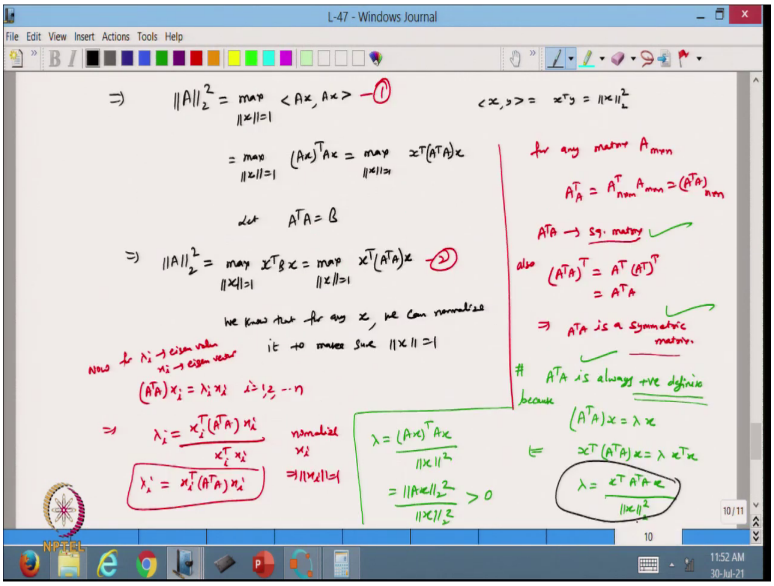
Lecture - 48
Sensitivity analysis of a system of linear equations

(Refer Slide Time: 00:17)



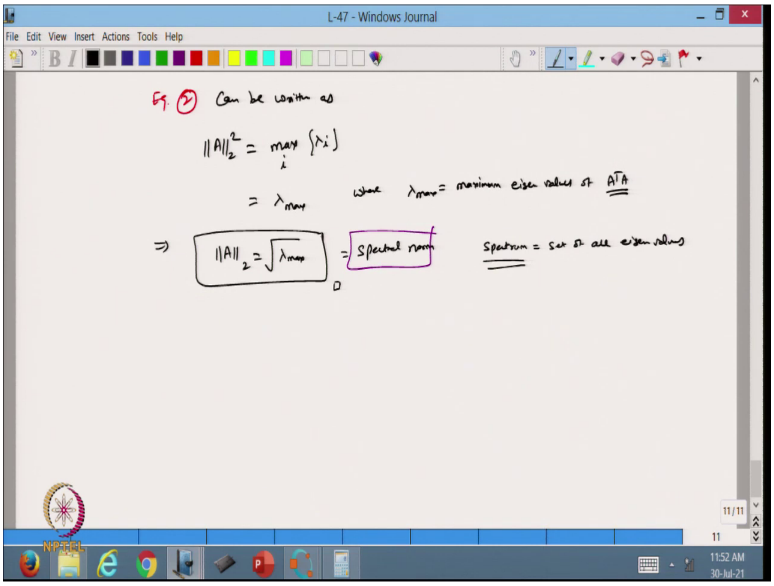
Hello viewers, so, welcome back to the course on Matrix Computation and its applications. So, in the previous lecture we have discussed the 2 norms of a given matrix. Now, in this lecture we will continue with that one. So, in the previous lecture we have discussed that for any matrix A, we have defined the 2 norm and that is equal to the square root of lambda maxima. Where lambda is the maximum, so, I can define that lambda maximum is the maximum eigenvalue of matrix A transpose A.

(Refer Slide Time: 01:25)



And, we have also seen in the previous lecture, that this matrix is positive definite and symmetric. Symmetric matrix all the eigenvalues will be real and positive definite means all the eigen values will be positive. So, that is why we are just taking the square root of this one.

(Refer Slide Time: 01:33)

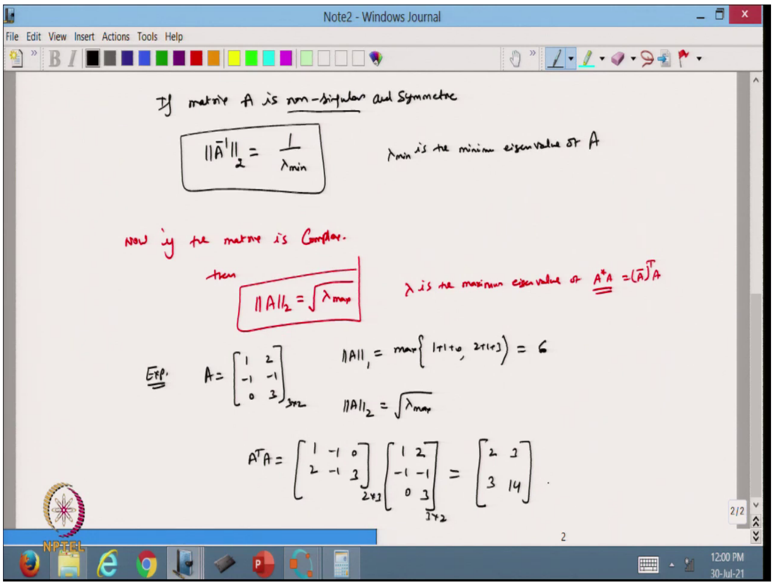


So, this norm is also called spectral norm, because we know that the spectrum we know is a set of all eigenvalues. So, then we are defining the maximum eigenvalue. So, this is called the spectral norm. So, these two norms are also called spectral norms.

Now, if A is a symmetric matrix, what is going to happen? Then, in this case $A^TA=A.A=A^2$. So, now, if I take the eigenvalue, the eigenvalues of A^TA is the same as the eigenvalues of A^2 and it means eigenvalues of A and then taking the square.

So, from here I can say that, now if λ is an eigenvalue of A, which implies λ^2 is the eigenvalue of A^2. So, from here I can define that, the spectral norm will become $\lambda_{\max}^2$ I can take from here and then it becomes maximum this one. So, now, in this case I can say that the 2 norm when the matrix is symmetric is equal to the spectral radius of matrix A. So, this way we can define the eigenvalue when the given matrix becomes the symmetric matrix.

(Refer Slide Time: 04:22)



Now, after this one now let us see that matrix A is non singular. Then, I can define the $\| A^{-1} \|_2 = \frac{1}{\lambda_{\min}}$. Where $\lambda_{\min}$ is the least or I can say this is the minimum value, minimum eigenvalue of because in this case transpose A ok.

So, I know that eigenvalues of A and it is inverse are or are inversely proportional to each other then from here I can define that, if A is a non singular matrix then this is equal to this one. So, I just assume that if matrix A is non singular and symmetric, then I can convert this one into A, because easily then we can define this value.

So, now, if the matrix is complex, then I can define the same way the $\| A \|_2 = \sqrt{\lambda_{\max}}$, where lambda is the maximum eigenvalue, maximum eigenvalue means the eigen value which has the maximum magnitude of A*A. All other things are same because A*A means A conjugate transpose A.

All other things will be the same and then we can define the 2 norms in the case of complex eigenvalue, complex eigen matrix or complex matrix. So, this way we can define this value.

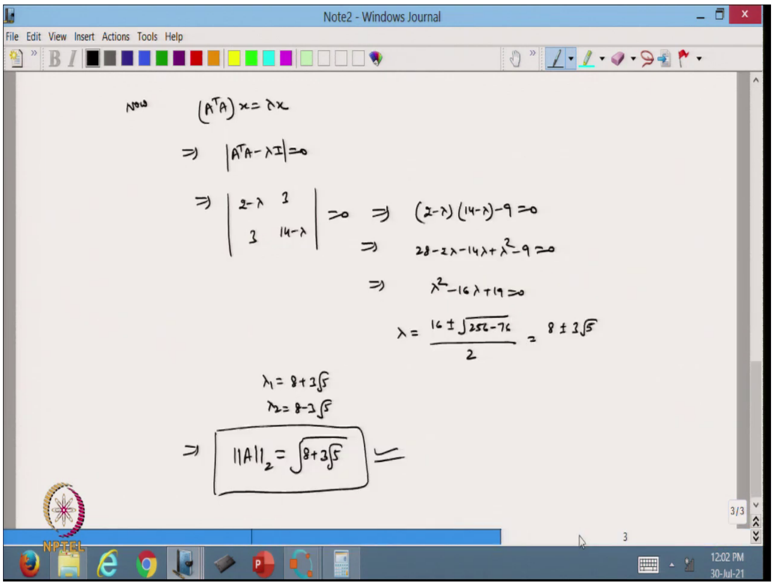
Now, suppose let us take one example, suppose I take the matrix $A = \begin{bmatrix} 1 & 2 \\ -1 & -1 \\ 0 & 3 \end{bmatrix}$. So, let us take this matrix. And, this is a rectangular matrix 3*2. Now, I suppose it is to define 1 norm. So, we already know that it is a maximum of the column.

So, I can define from here so, $\| A \|_1 = \max\{1+1+0, 2+1+3\} = 6$. So, everything we already know, now I can define it as 2 norms. Now, $\| A \|_2 = \sqrt{\lambda_{\max}}$, I need to find $\lambda_{\max}$ of this one.

Now, first try to define what is $A^T A$? So, $A^T A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & -1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 14 \end{bmatrix}$

So, I got this matrix that is $A^T A$. And, of course, this is A symmetric matrix that we can define from here.

(Refer Slide Time: 09:27)



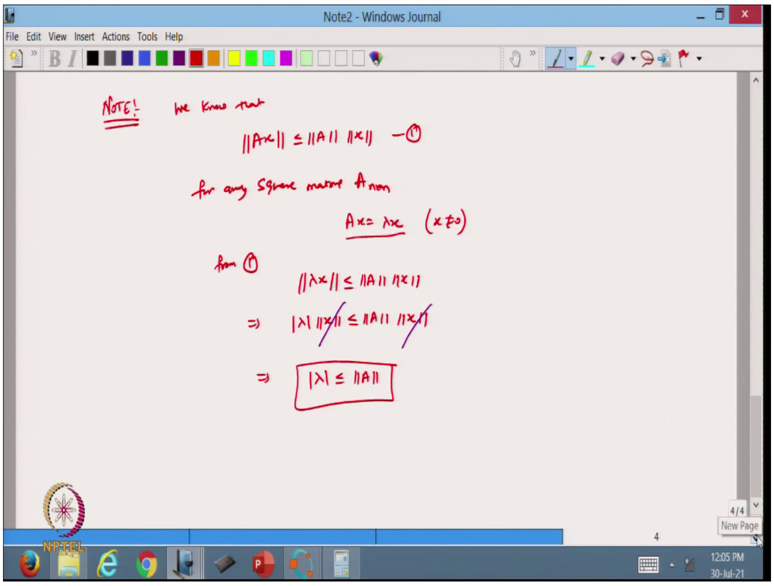
$$\left(A^T A\right)x = \lambda x$$

Now, I want to define it is eigenvalues. So, from here I can define $|A^T A - \lambda I| = 0$

$$\begin{vmatrix} 2-\lambda & 3 \\ 3 & 14-\lambda \end{vmatrix} = 0$$
$$(2-\lambda)(14-\lambda) - 9 = 0$$
$$28 - 2\lambda - 14\lambda + \lambda^2 - 9 = 0$$
$$\lambda^2 - 16\lambda + 19 = 0$$
$$\lambda = \frac{16 \pm \sqrt{256-76}}{2} = 8 \pm 3\sqrt{5}$$
$$\lambda_1 = 8 + 3\sqrt{5}, \lambda_2 = 8 - 3\sqrt{5}$$

So, I got these eigenvalues. So, from here I can say that, the 2 norm will be $\|A\|_2 = \sqrt{8+3\sqrt{5}}$ So, this is my spectrum of the matrix A. So, here we are taking the eigenvalue of A transpose A and then we got this value. So, this is my spectral norm of the given matrix. Now, after doing this one, we can have a relation between the eigenvalues and the norm. So, from here I just can make a note.

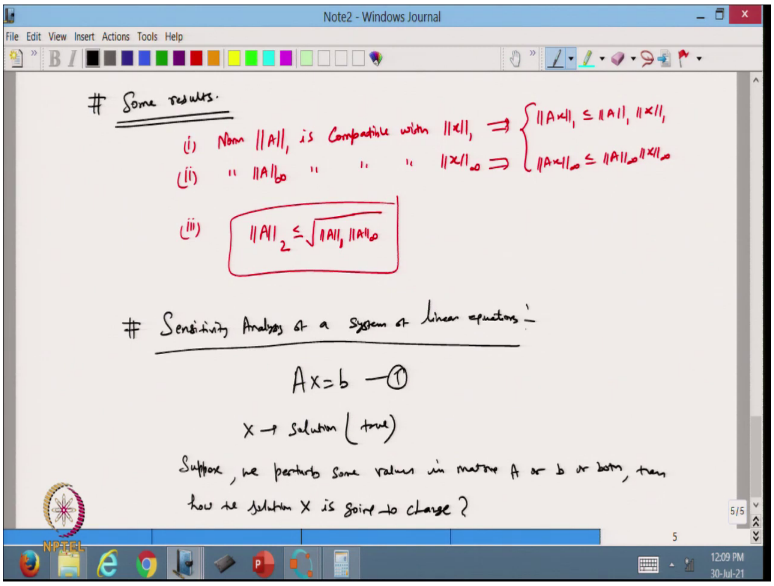
(Refer Slide Time: 12:05)



We know that, that $||Ax|| \leq ||A||.||x||$ I can write like this one. So, it is a vector norm and matrix norm compatibility condition. Now, for any square matrix A, so that is A square matrix we can define $Ax = \lambda x$. So, from one I can write it is $||\lambda x|| \leq ||A||.||x||$ and I know that $x \neq 0$ Only then can we find the eigenvalues. So, from here this become by the properties of the norm this become $|\lambda| ||x|| \leq ||A||.||x||$

And, now $x \neq 0$. So, from here I can cancel out this one and from here I can say that this condition. So, it says that for a matrix A, if I take it is eigenvalues and taking the magnitude, then $|\lambda| \leq ||A||$. So, this gives the upper bound for the matrix. For the eigenvalues of a given matrix, the magnitude of the given matrix, magnitude of the eigenvalue. So, this is one of the important relations we have discussed.

(Refer Slide Time: 14:27)



Now, from here on we are going to define a very important concept and before that I will just write some results. So, in this case we are going to take some results, 1st one is that $\| A \|_1$ is compatible with vector norm $\| x \|_1$. the 2nd one is that the infinity norm is compatible with the infinity norm. So, basically in this case $\| Ax \|_\infty \le \| A \|_\infty . \| x \|_\infty$, $\| Ax \|_1 \le \| A \|_1 . \| x \|_1$

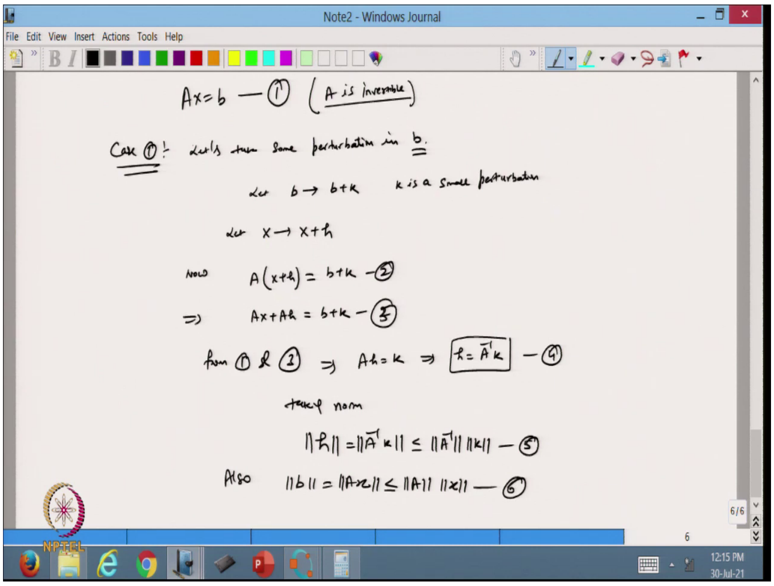
So, that is the compatibility condition. So, after doing this one more thing we want to discuss is 3rd property, that if you take the $\| A \|_2 \le \sqrt{\| A \|_1 \| A \|_\infty}$

Because, we have seen that 1 norm was the biggest 1 and 2 norm was in between and infinity is the smallest 1 in that case. So, I can say that 2 norms always satisfy this condition. And, this is just you can see that it is A type of geometric mean of norm 1 and infinity.

So, this is also one of the conditions we can keep in mind. So, after doing these things we are going to start an important topic that is called the sensitivity analysis of a system of linear equations. Sensitivity means suppose I have a system Ax = b. So, this is a system I have, now given the system there, then x is its solution, I can say it is a true solution, I solved it and I got this value.

Now, suppose we perturb some values in matrix A or b or both, then, how then, because when we perturb some values in the matrix A or b or in both, then how the solution x is going to change? So, this one we want to see, what is going to happen, when we do the small change in the given matrix or on the right hand side, because why are we doing this one?

(Refer Slide Time: 19:11)



Because, we know that, when we have a system Ax =b and we want to solve this system, this system may be very big and we want to solve this one. So, we have to take the help of a computer. And, in the computer we have different types of error, round of error, or discretization error that in that case sometimes the small change in the given right hand side vector or in the matrix, can lead this system to an infinite number of solutions or no solution.

So, that is why we have to check the sensitivity of the given system with respect to the perturbation in the values of A or b. So, this one we just want to study. So, let us take this Ax = b. And, now suppose I take the concept that case 1, what are we going to do is that?

Let us take some perturbation in b, b means right hand side vector. Now, let b change to b +k where; k is a small perturbation. Perturbation means small value. And, let x will change to x +h let us see. Now, I want to see how big the h will be ok. So, this one I want to see. Now, A(x +h)=b+k

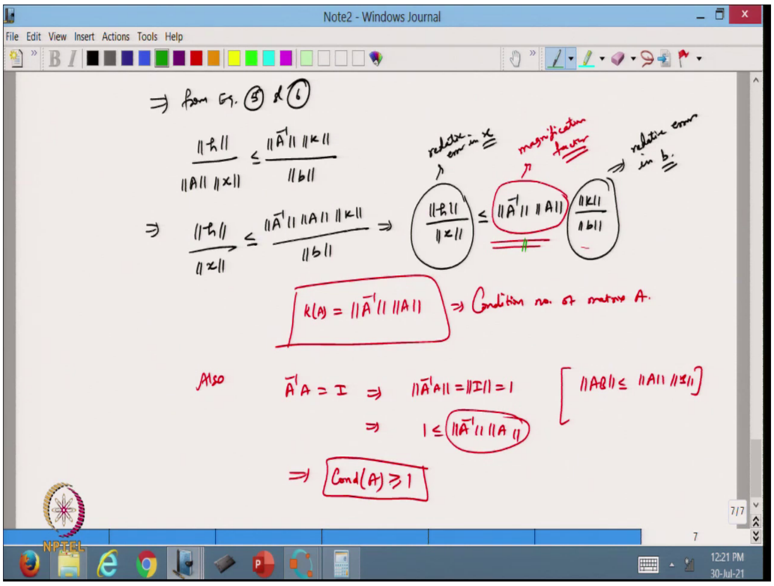
➪ Ax+Ah=b+k

So, I call this as 2. And, here in this case I am considering that A is invertible, because I am considering that this has a solution that is unique, so A is invertible. Now, 2 minus 1 or from 2 and 1, I can write.

So, I should write it is 3. So, from 1 and 3, I can write from here. So, in this case I can write because from 3 I subtract 1 then I will get Ah = k, this value I am going to get. And, from here I can write my $\| h \| = \| A^{-1}k \| \leq \| A^{-1} \| . \| k \|$

$\| b \| = \| Ax \| \leq \| A \| . \| x \|$

(Refer Slide Time: 24:09)



So, this quantity Ax is bigger than this one so from equation 5 and 6. Now, my h norm is less than this and this is the bigger one. So, I can write from here my h divided by suppose I divide by this one. I can write as A inverse k and divided by b. Because, I am dividing so, this quantity is bigger by than b. So, if I divide by this one, then this sign will be there.

Now, after doing this one, we can write from here, I can write h and k, this one, I can write A inverse norm, A norm and then x, b or maybe I should take this one as ok. So, let us keep it x here and let us keep it k here. Now, from here I can write the norm of h divided by norm of x, that is less than equal to A inverse norm, A norm and this is I can write k by this one.

$$\frac{||h||}{||A||.||x||} \leq \frac{||A^{-1}||.||k||}{||b||}$$
$$\frac{||h||}{||x||} \leq \frac{||A^{-1}||.||A||.||k||}{||b||}$$

Now, if you see from here, then this is the relative error in b, because we have changed the b from b to b + k. So, it is the relative error in the b with respect to b. So, that is a relative error. This is the relative error in x. Now, after doing this one, we found a new thing here. Because, this is the magnification factor, because suppose this quantity is there, whatever this quantity is there.

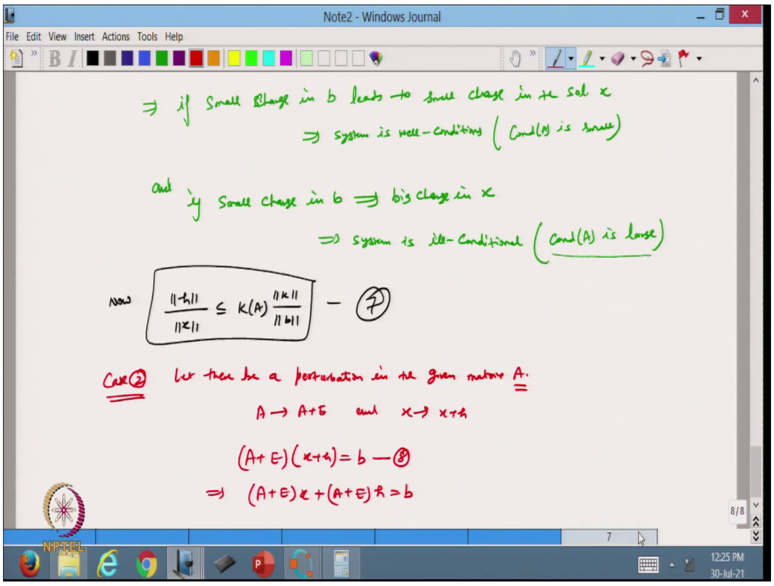
If this value is very large, then it may happen that this quantity is going to be very large. And, then I can say that a small relative change in b leads to a very big change in x. If this quantity is very large, but if this quantity is small 1, then we can say that a small change in the right hand side leads to a small change in the value of the solution.

So, everything depends upon this factor. So, from here I take this factor as A inverse norm into norm A and I define this as k A. So, this factor is called the condition number of matrix A. So, this is called the condition number of the matrix A. Only condition is that A should be invertible. Also, I know that A inverse A is equal to I. And, from here I can write A inverse A taking the norm both sides are equal to I and norm of I is always 1, then from here.

At this quantity I can take the condition so, 1 is there. So, that is always less than the norm of A into norm. I am sorry, I am writing here about this condition. So, from here I can say and this is what is the condition number. So, I can say from here that the condition number of the matrix A is always greater than equal to 1, so, now, if this quantity is becoming equal to 1.

So, now, if this is equal to 1, then small change in this case will lead to the small change in here. Because, if it is 1, then even the small change here the small change in x will be smaller than the change in here. So, in that case the system will be a condition. So, now, from here I see that the condition number of the matrix is always greater than equal to 1.

(Refer Slide Time: 29:36)



So, from here I can say that, if small change in b leads to small change in the solution x, implies that the system is well conditioned. And, in that case the condition number of the matrix is small , so it is small. And, if small change in b implies big change in x so, we say that the system is ill-conditioned. And, this is possible when the condition number of the matrix A is large.

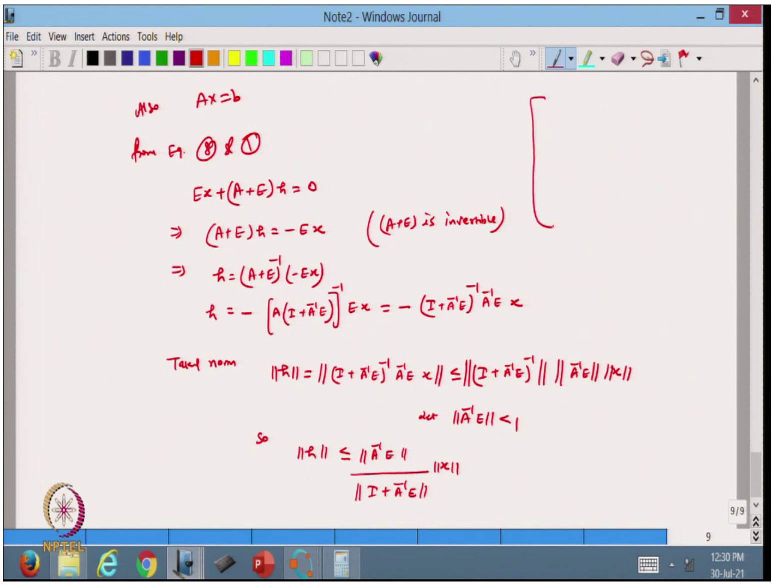
So, large means when the condition number of the matrix is close to 1, then it is well conditioned. Otherwise, it is ill-conditioned because the small and the large are relative terms. So, when this is close to 1, then it is called the condition number is good enough. And, if it is a very big number, then it is going to be called an ill conditioned matrix or it is going to have a large condition number of the given matrix.

So, this way we can discuss the condition number of A given matrix. Now, once I get this value condition number. So, from here now we can write the relative change. So, relative change is here. So, the relative change in x is always less than equal to the condition number of the matrix A, and then this is the relative change that has happened in b.

So, this is one of the terms we are going to use. So, that is my equation number 7. So, everything depends on the condition number. Now, from here so this is the case 1 we have taken; now I am defining the case 2. So, let there be a perturbation in the given matrix A.

So, let us take the perturbation in the matrix A. So, suppose my matrix A changes to A plus E small perturbation there. And, my x changes to again x plus h. Now, I can write from here A plus E and my solution becomes x plus h and that is equal to my b ok. So, let us write it as equation number 8, in this case. From here I can write this as (A + E)(x+ h) =b.

(Refer Slide Time: 34:05)



Now, also Ax =b, this is the equation number 1 we have taken. So, I can write now from equation 8 and 1. So, I can write from here Ax =b and b will cancel out. So, from here I can write that I get my Ex +(A + E)h=0

And, from here I can write that (A +E) h =-Ex. Now, my matrix I am considering is that A is invertible and A+ E is also invertible. So, A plus E is invertible, because only then we are able to find the solution. So, a solution exists there. So, I can write h= $(A+E)^{-1}(-Ex)$

$$= -[A(I + A^{-1}E)]^{-1} Ex = -[(I + A^{-1}E)]^{-1} A^{-1} Ex$$

Now, I am taking the norm. I can define the norm of h as so the norm of this one is just I am writing negative signs will disappear.

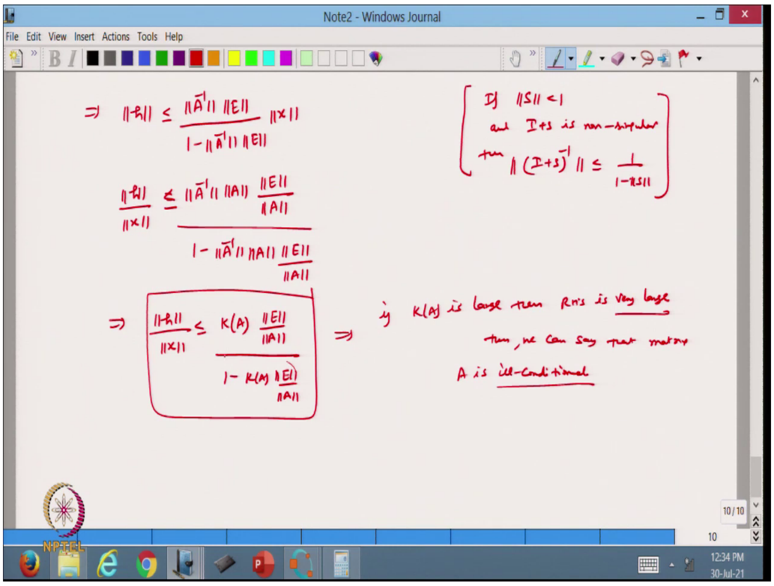
$$\| h \| = \| [(I + A^{-1}E)]^{-1} A^{-1}Ex \|$$

So, this will become $\leq \| [(I + A^{-1}E)]^{-1} \| . \| A^{-1}E \| . \| x \|$

And, now from here I can define now let. So, I define here that let norm of A inverse E is less than equal to 1.

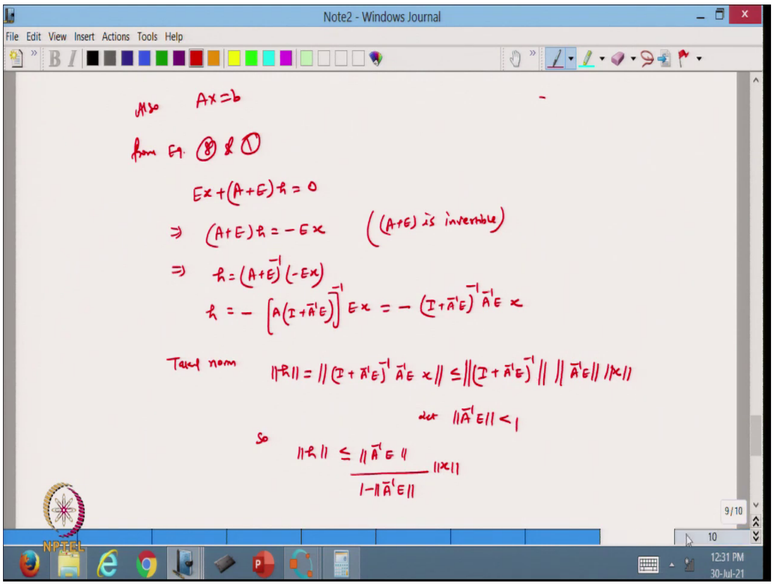
So, let us take this one. Because, E is say small change in the A so, we consider that $\| A^{-1}E \|$ <1. So, I can define my $\| h \| \leq \dfrac{| A^{-1}E \| . \| x \|}{(I + A^{-1}E)}$

(Refer Slide Time: 38:30)



This is what we are going to use, if suppose S is A matrix and norm of S is less than 1 and I plus S is non singular, then I plus S inverse norm is always less than 1 over, 1 minus this one. So, in fact, I am not going to use this concept here directly. So, I can write from here, then it should be equal to taking the same way.

(Refer Slide Time: 39:18)



So, I can write this as a 1 minus A inverse E norm, this one into x. So, this one we have $\| h \| = \| [(I + A^{-1}E)]^{-1} A^{-1}Ex \|$ defined. So, from here I can write that norm of $\le \| [(I + A^{-1}E)]^{-1} \| . \| A^{-1}E \| . \| x \|$ Now, it can be written as A inverse I can write here this and this one I can write, this

$$\| h \| \le \frac{| A^{-1}E \| . \| x \|}{(I + A^{-1}E)}$$

$$\le \frac{| A^{-1}E \| . \| x \|}{1 - \| A^{-1}E \|}$$

$$\le \frac{| A^{-1} \| . \| E \| . \| x \|}{1 - \| A^{-1} \| . \| E \|}$$

$$\frac{\| h \|}{\| x \|} = \frac{| A^{-1} \| . \| A \| . \frac{\| E \|}{\| A \|}}{1 - \| A^{-1} \| . \| A \| . \frac{\| E \|}{\| A \|}}$$

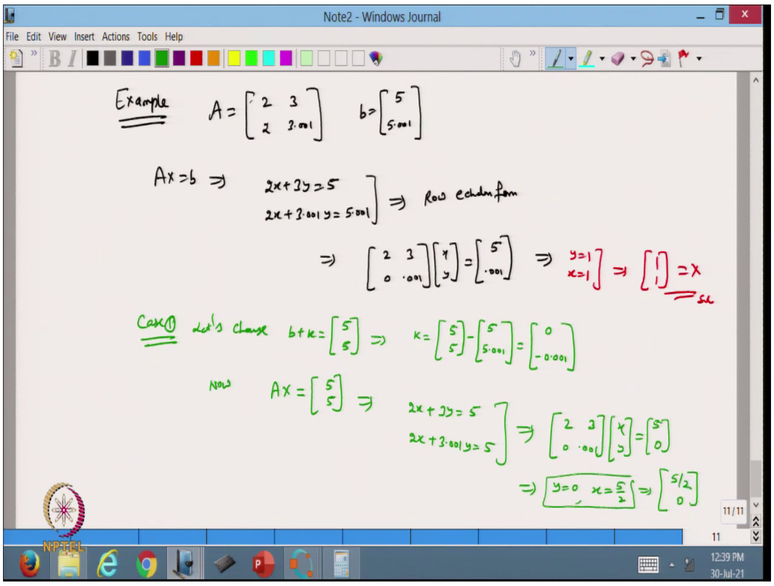
$$\frac{\| h \|}{\| x \|} = \frac{K(A) . \frac{\| E \|}{\| A \|}}{1 - K(A) . \frac{\| E \|}{\| A \|}}$$

and this is again the relative change in A. So, in this case also if the relative change in A this is small, but the condition number is very large.

So, if the condition number is very large this quantity is going to be very large and 1 minus 1 this is going to be very small. So, then the whole quantity will become very large. So, it implies that if the condition number of the matrix is large, then the right hand side is very large.

Then, we can say that matrix A is ill conditioned. And, the system is also ill conditioned. So, everything depends upon the condition number here and that is the condition we can define, when we take the sensitivity analysis of the given matrix. So, let us take one example and let us see that, what is the meaning of this?

(Refer Slide Time: 42:36)



Suppose, I take A matrix A let us take a simple matrix $A=\begin{bmatrix}2 & 3\\ 2 & 3.001\end{bmatrix}, b=\begin{bmatrix}5\\ 5.001\end{bmatrix}$. Let us take this one. Now, Ax =b which implies that, it should be 2 x + 3y = 5 and 2 x +3.001 y =5.001, suppose I take this one.

Now, if we convert this into the echelon form. So, I convert this row echelon form, that gives me $\begin{bmatrix} 2 & 3 \\ 2 & 3.001 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 5.001 \end{bmatrix}$ And, if you solve this condition system then you will get

$$y = 1, x = 1$$
$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = X$$

That is my x solution ok. Now, what is going to happen now? I will let us take case 1.

So, let us change b, a new b as so, whatever the b is there, I change maybe I can define here

$$b + K = \begin{bmatrix} 5 \\ 5 \end{bmatrix} => K = \begin{bmatrix} 5 \\ 5 \end{bmatrix} - \begin{bmatrix} 5 \\ 5.001 \end{bmatrix}$$

$$K = \begin{bmatrix} 0 \\ -0.001 \end{bmatrix}$$

So, this is what we have defined. So, K is this one.

So, that is the small change we have done on the right hand side b. Now, let us see what is going to happen? So, we are going to solve this condition now so, I have a

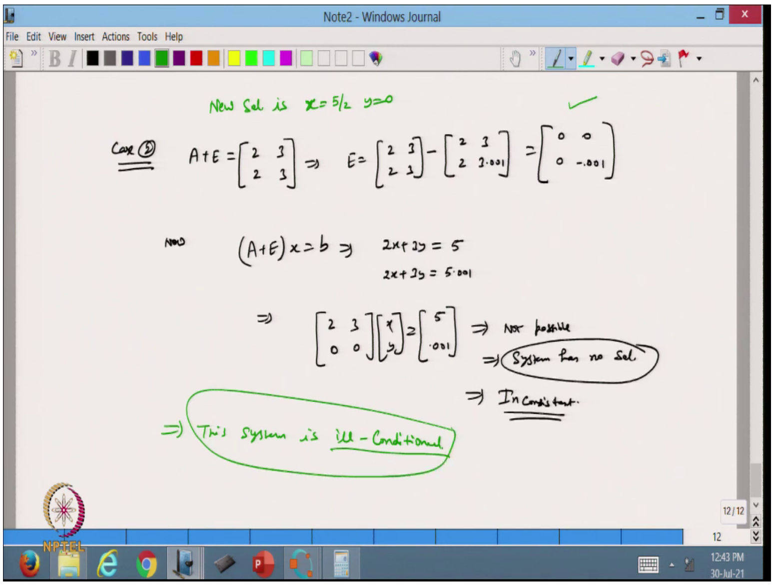
$$AX = \begin{bmatrix} 5 \\ 5 \end{bmatrix} => 2x + 3y = 5, 2x + 3.0001y = 5$$

Now from here, if you see I can write in echelon form.

$$\begin{bmatrix} 2 & 3 \\ 0 & 0.001 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$
$$y = 0, x = 5/2$$

$\begin{bmatrix} 5/2 \\ 0 \end{bmatrix}$ is the solution.

(Refer Slide Time: 47:08)



So, now, my new solution is $\begin{bmatrix} 5/2 \\ 0 \end{bmatrix}$. So, you can see from here that a small change on the right hand side can lead to a big change in the solution . So, this is what we got from here. Also, let us take another case, let us take another case, case 2.

$$A+E=\begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix} => E=\begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}-\begin{bmatrix} 2 & 3 \\ 2 & 3.001 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & -0.001 \end{bmatrix}$$

Now, I take the matrix A + E the change the matrix and let us take this matrix as

so, this is what I have taken. So, it is a very small change in this one.

So, my E is this one. Now, I want to see what is going to happen in the solution, let us see. Now, it is

$$2x+3y=5, 2x+3y=5.001$$

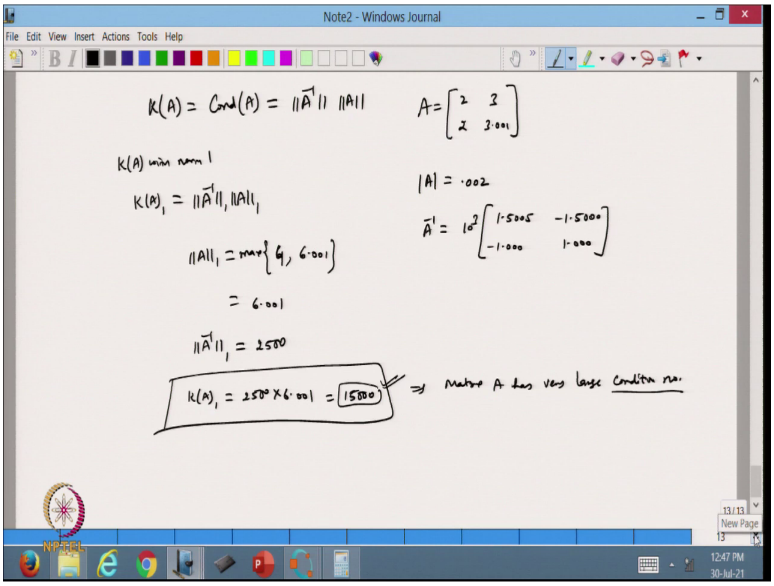
So, we have taken this value. Now, from here if you see, if I write this in the echelon form. So, you will get

$$\begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 0.001 \end{bmatrix}$$

So, in this case I can say that in the last equation, we can say that it is not possible. So, the system has no solution or I can say that inconsistent the system is inconsistent there is no solution. So, you can see from here a small change in the given matrix leads to the no solution of the given system. So, from these things I can conclude that this system is ill-conditioned. I have not seen it till now we have not seen, what is the condition number?

But, looking at this behavior I can say that this system is a condition. Because a very small change you can see in the given matrix leads to the system having no solution, over a very small change in b leads to the very big change in the given solution. So, from here I can conclude that the system is ill-conditioned. Now, let us find out the condition number of this matrix.

(Refer Slide Time: 51:10)



So, let us see what is going to happen for this matrix? Or I can call it the condition number of a given matrix. So, this is going to be norm of A inverse and A, where my $A = \begin{bmatrix} 2 & 3 \\ 2 & 3.001 \end{bmatrix}$.

Now, quickly we can find out this value. So, now, I can say that, my A matrix is this one and if you see the determinant. So, the determinant of this matrix it will be 0.002 and it is inverse if you find out the inverse is coming

$$A^{-1} = 10^{3}\begin{bmatrix} 1.5005 & -1.5000 \\ -1.000 & 1.000 \end{bmatrix}$$

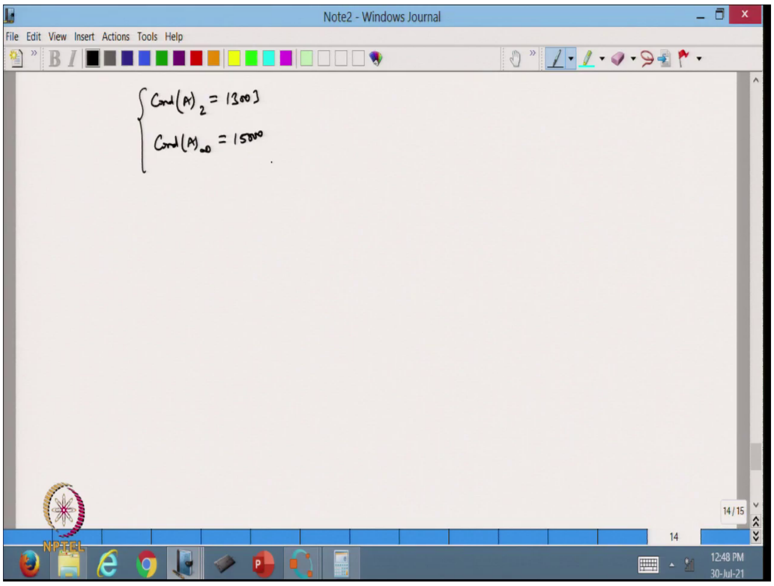
So, this is the inverse we have already calculated and its value is this one. Now, I can find out the condition number. So, I can find out the condition number with any of the matrix norms. So, let us find out the condition number of the matrix A with norm 1.

So, I can write as A with norm 1, it means I am taking 1. So, from here if you see then I can find out the norm 1 of the matrix. So, it is taking the maximum columns. So, it is maximum$\{4, 6.001\}$. So, its value is 6.001. Now, A inverse 1 norm I am taking. So, if we calculate then it is going to be the maximum will be 2.5 and here it will be 2.5 multiplied by 3.

So, the value is coming to 2500. Now, from here if you see the condition number K(A) in the norm 1 that is going to be 2500 X 6.001 that is approximately equal to 15,000. So, you can see that the condition number of the matrix is very very large; we expect that the condition number of the matrix should be close to 1, but here it is very very big.

So, from here I can say that this matrix A has a very large condition number. So, from here we can say that this system is ill-conditioned. Now, the same way we can define the condition number in another format also. So, in the computer we have seen how we can find the condition number of a given matrix?

(Refer Slide Time: 55:08)



I have seen that the condition number of the matrix A in 2 norms was coming from this value. And, condition number of A in the infinity norm we have seen, that this is coming 15,000. So, based on this one we because we know that all norms are equivalent. So, in this case we have seen that the condition number of the matrix is very very large.

And, therefore, I can say that the given system, that is Ax equal to b is very sensitive to the change and hence it is a very ill conditioned system. So, we will stop here. So, in the lecture today we discussed the condition number of the matrix.

And, we have seen how the condition number of a matrix is going to be used to study the sensitivity analysis of the given system of equations that is Ax equal to b. So, in the coming lecture also we are going to discuss the condition number of matrices and other concepts. So, thanks for watching.

Thanks very much.