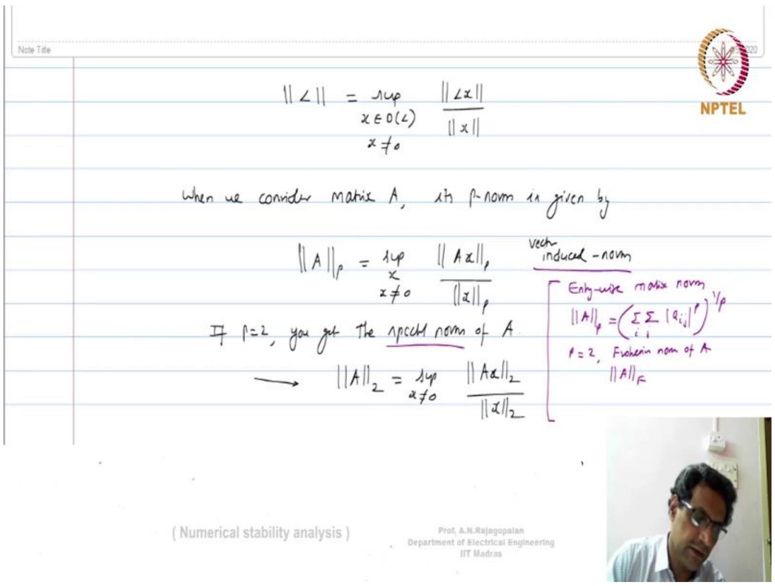
Image Signal Processing
Professor A. N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture 5
Numerical Stability Analysis

(Refer Slide Time: 00:21)



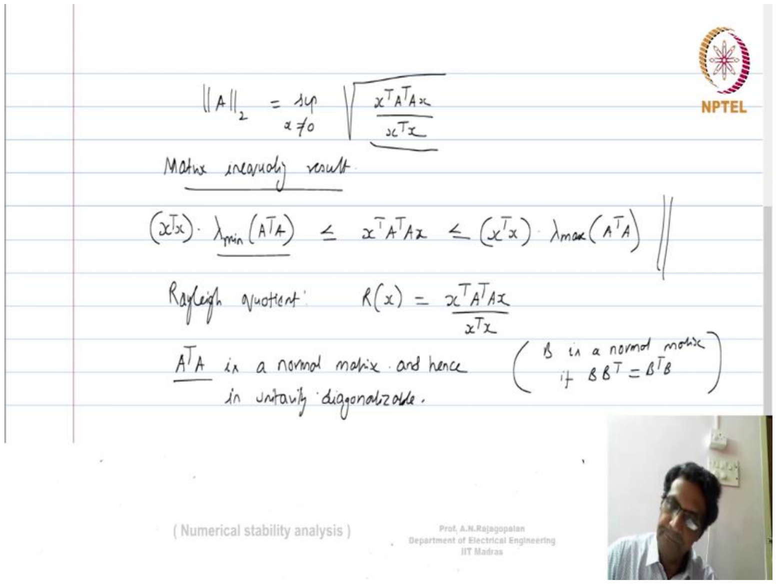
So, we saw that the norm of linear operator is given as norm l is equal to sup over x belonging to domain of l, it is not equal to 0, norm l x by norm x and now the idea is that now although for a matrix norm, so when we consider the special case of a matrix when we consider matrix A, then it is p norm is given by, p norm is given by norm A p is equal to sup over 2x not equal to 0, sup of x is not equal to 0 and norm Ax, p norm by norm x is also p norm, but the idea is that, now such a norm, I mean we have already, we have seen other kinds of norms, we have seen the Frobenius and all of that and that norm and which I incidentally was also for, was also another matrix norm.

Now, if you compare that matrix on with this norm this is the norm that comes via a form like this is called an induced norm or it is called a vector induced norm because it is actually induced by this vector, norm Ax p by norm xp and that gives you p norm, for example, if p is equal to 2, p is equals to 2 you get what is called the spectral norm of A for p equal to 2 you get the spectral norm of A which is what, which is what is going to be of interest to us, this is because the spectral norm a certain has special properties that can actually that can help us understand the numerical stability of a matrix and therefore the spectral norm is given as norm 2 is equal to sup, it is not equal to 0, the norm Ax, to norm of Ax, the vector Ax and 2 norm of the vector x.

Now, compare if you compare this with the with the entry wise norm, the other norm, now this is anywhere it just as an aside okay for those of you who might be this interested in knowing this is a vector induced now, right. Whereas on the other hand if you have entry wise norm which is what we have seen earlier when we did transforms, all the image transforms and all we had what we called entry wise matrix norm and that norm was actually given by norm A p is equal to if you recollect we had it like summation overall entries of A and i, j and then magnitude Aij, p the whole power 1 by p, this is called an entry wise norm and p equal to 2 was actually a Frobenius norm, p equal to 2 gave us the Frobenius norm of A, so which we denoted as norm A and then we put F here that was the Frobenius norm.

Now, p equal to 2 now, for p equals to 2 here you get this form and this is called spectral norm. So, a matrix has various different norms it is also something called the Schatten norm and so on which is a obviously singular values and so on, but right now our focus is on this spectral norm. It is called the spectral norm because it eventually turns out that this norm can be explicit in terms of the Eigen values of this matrix, therefore it is called the spectral norm. We will see all of that as we move along.

(Refer Slide Time: 04:05)

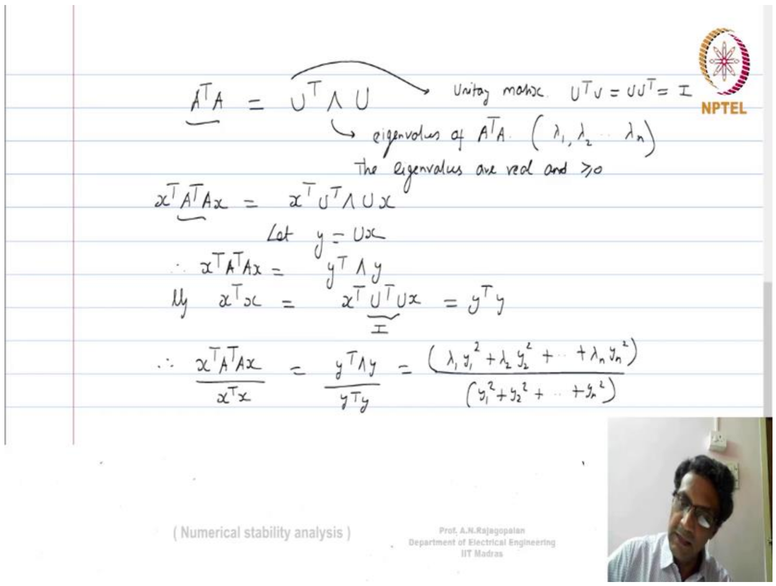


Now let us get back to this expression on A the spectral norm is equal to sup over x not equal to 0, now I can write norm Ax by norm x root of x transpose A transpose Ax upon x transpose x, turns out that it was the same as. Now, there is a matrix inequality that says there is a matrix inequality, inequality result which says that equality result, which says that lambda x transpose x scalar times lambda min of A transpose A, so this is the minimum Eigen value A transpose A is less than or equal to x transpose A transpose Ax and this in turn is less than or equal to x transpose x into lambda max of A transpose A, where x is a vector and so on.

Now, it is this inequality that will try to use here when we want to compute the spectral norm of A but then we can actually show this result, we can actually show this inequality and to show that let us kind of look at what is called a Rayleigh quotient of the matrix, now Rayleigh quotient is given as R of x is equal to x transpose A transpose Ax upon x transpose x in fact the term sitting here is in fact this is is a Rayleigh quotient.

Now, if A transpose A is actually a normal matrix, correct, normal matrix now what is a normal matrix? We have already seen while doing image transforms, B is a normal matrix if B transpose B B B transpose is equal to B transpose B and of course, when the B is symmetric this is automatically obey, A transpose A is the symmetric matrix and therefore it does obey this and therefore it is a normal matrix and what is so special about a normal matrix? A normal matrix is unitarily diagonalizable, and hence is unitarily diagonalizable this again something that we have seen when we did image transform, unitarily diagonalizable, unitarily diagonalizable.

(Refer Slide Time: 06:48)

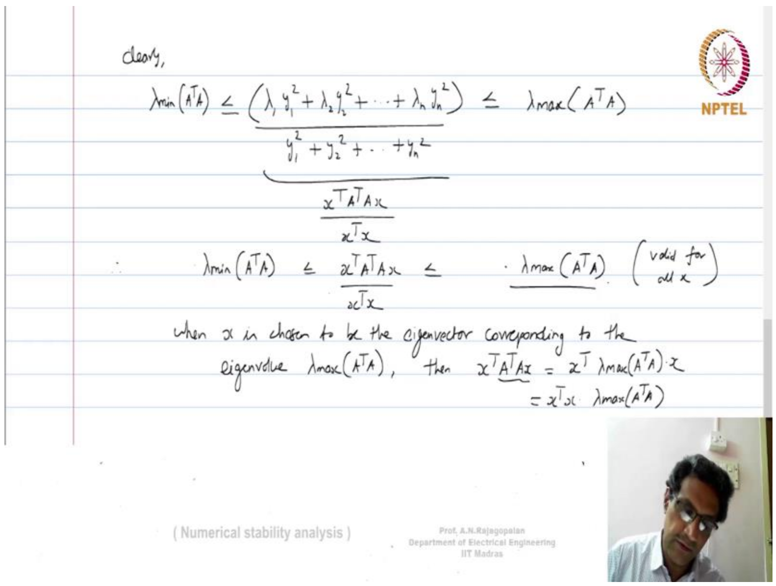


Now, what does that mean? That means that A transpose A can be written in terms of the unitary transform as Y transpose diagonal U, that is diagonal contains Eigen values remember A transpose A, not A, Eigen value of A transpose A and because of the fact that now A transpose A is symmetric, therefore the Eigen values are real and greater than or equal to 0, is a real and greater than or equal to 0, because of the fact that this is the symmetric matrix.

Now, what this means is that I can rewrite x transpose A transpose Ax as x transpose then I replace A transpose A by U transpose diagonal U and then I get back to x. Let us say y be equal to Ux, therefore x transpose A transpose Ax has been written in terms of the unitary transform becomes y transpose diagonal y. Similarly, x transpose x can then be written as x transpose U transpose Uy because U transpose U is identity it is a unitary transform, we remember this is a unitary transform, the unitary matrix that is U transpose U is equal to mu U transpose is equal to identity.

Now, therefore this then becomes equal to y transpose y if I simply put substitute for oh sorry this is x, substitute in for Ux is y then I get y transpose y, therefore x transpose A transpose Ax for x transpose x becomes equal to y transpose diagonals y by y transpose y, but if you could call the Eigen values, if these Eigen values are lambda 1, lambda 2 all the way up to lambda n, then what we have is this becomes equal to lambda 1 y1 square plus lambda 2 going to square all the way up to plus plus plus lambda n y n square upon y1 square plus y2 square plus plus plus yn square.

(Refer Slide Time: 09:21)



Now, this ratio clearly, lambda 1 y1 square plus lambda 2 y2 square plus plus plus lambda n yn square upon y1 square, now this clearly is less than or equal to divided by let us say let us divided by y1 square plus y2 square plus plus plus yn sqare this ratio clearly, let me write this clearly at top because I will try to use something for the left, so clearly if this should be less than or equal to lambda max of A transpose, of A transpose A because each is an Eigen value of A transpose A, this sum, this ratio in fact it should be less than or equal to n should be greater than or equal to lambda min of A transpose A, it is obvious because this y2 square y2 square, because we can take lambda max as, lambda max out then of course, actually you do not need the rest are all less than lambda max and therefore this should be less than or equal lambda max of A transpose A.

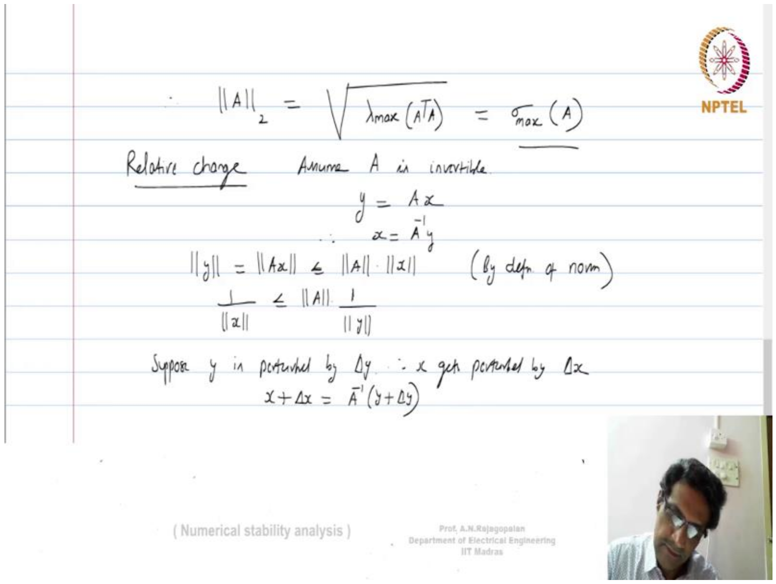
Similarly, if you try to pull lambda min then the rest are all greater than lambda min therefore this ratio should be greater or equals to lambda min of A transpose A. Now this ratio is nothing but as we saw earlier this can be rewritten as x transpose A transpose Ax, this is nothing but x transpose x written in terms of x and therefore, that result that I wrote down follows x transpose x into lambda min A transpose A is less than or equal to x transpose A transpose Ax because this x transpose x goes and multiplies on the left and the right is equal to, less than or equal to x transpose x into lambda max of A transpose A.

Now, this is an upper bound, now this upper bound can even be met, the upper bound is met when because we are interested in the supremum, the upper bound is met, now the same way we can also write this as, we can just divide this whole thing by x transpose x and therefore we can write this as upon x transpose x, so this goes away from here, this just let me remove the underline thing and then let me remove this from here, this is less than equal to lambda max.

Therefore, so such kind of back to the same equation here and this is the ratio for sup we are trying to find out and this will equal this, when x, okay this is valid for all x, this is valid for all x, for all x when x is chosen to be the Eigen value for sorry, it is chosen to be the Eigen vector corresponding to chosen to be the Eigen vector corresponding to the Eigen value, Eigen value lambda max of A transpose A, if you choose because it is valid for all x therefore let me choose x to be the Eigen vector corresponding to the maximum Eigen value that is lambda max of A transpose A.

Then x transpose A transpose Ax becomes equal to x transpose A transpose Ax it becomes then equal to lambda max of A transpose A into x and this x transpose which is equal to x transpose x into lambda max of A transpose A and this x transpose x will knock off this x transpose x and you get lambda max of A transpose A as an equality.

(Refer Slide Time: 13:35)



And therefore now the spectral norm of A is given as root of lambda max of A transpose A and in terms of singular values we know that the positive square root of the Eigen value of A transpose A is nothing but the singular value of A therefore this is equal to sigma max, so back to your singular value things, the single value, the singular decomposition seems to be coming up every now and then, it has a very very major role to play.

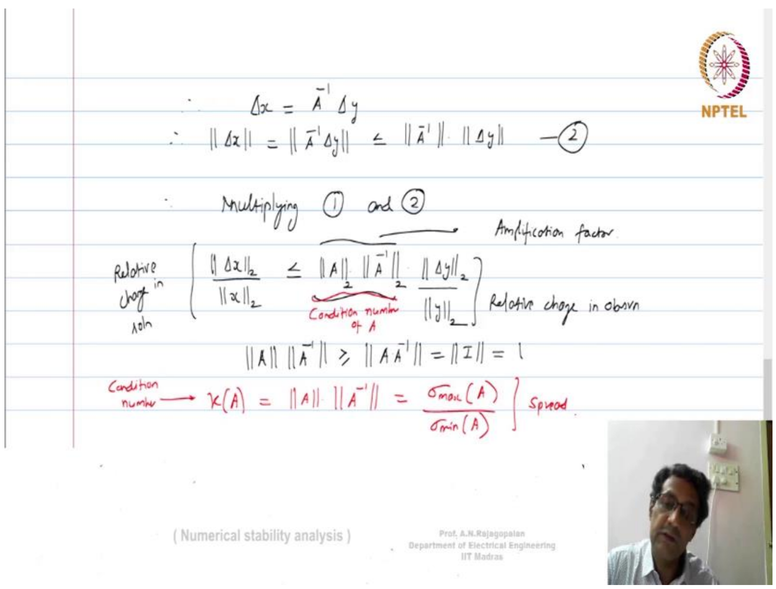
Now, having said this so we have these the spectral norm of matrix A, therefore the spectral norm of a matrix A is simply the maximum singular value of A. Now what is of interest is this notion of numerical stability, we have been trying to argue as to what will be the change in the solution, will we have to change in the solution relative to the change the observation, so there is a little change for example, let us now in the light of what we know about a matrix spectral norm, let us examined relative change.

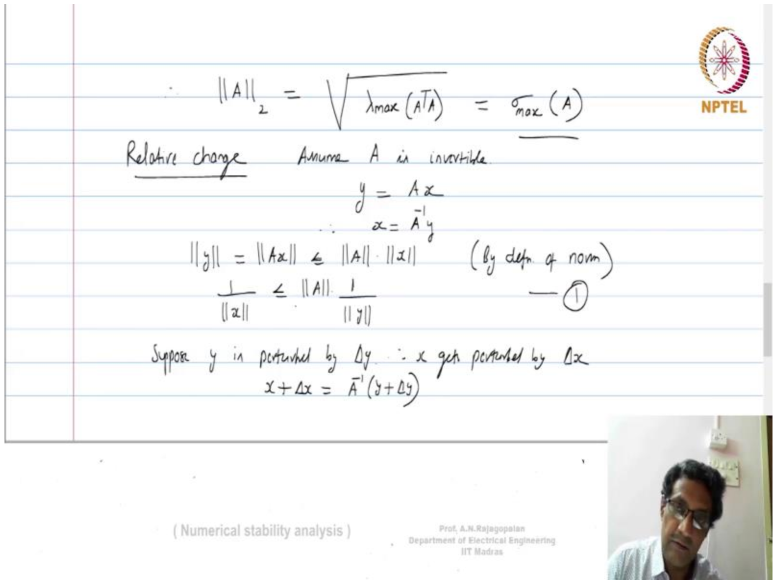
Suppose we have y is equal to Ax and assuming, A is invertible which is the most ideal situation that you can have, if it is not invertible then anyway we cannot do much but suppose, A is invertible therefore what do we know? We know that x equal to A inverse x we know this.

Now, let us look at norm of y it is equal to norm of Ax that is equal to norm of, that is less than or equal to we know is norm of A into norm of x, this is by definition of norm of the matrix norm. Now or in other words what you can say is 1 by norm x, 1 by norm x is less than or equal to norm A into 1 by norm y.

Now, x equal A inverse y this is not x into A inverse x this is x inverse y. Suppose y is perturbed like a small value by delta y then, because of the fact that A is invertible, for x perturbed by delta x or in other words when y is y plus delta y x becomes x plus delta x, x plus delta x is equal to A inverse y plus delta y or in other words or from this it follows that delta x follows that.

(Refer Slide Time: 16:42)





Therefore delta x is equal to A inverse delta y for norm delta x equal to norm A inverse delta y less than or equal to norm A inverse norm delta y, let us called as equation two, let us call the earlier one equation one, if multiplied 1 and 2, now if you multiply 1 and 2 because of the fact that every term here is a number greater than or equal to 0, I have only norms are involved and therefore even if you multiply the left and the right hand side quantity is the inequality rate it does not flip.

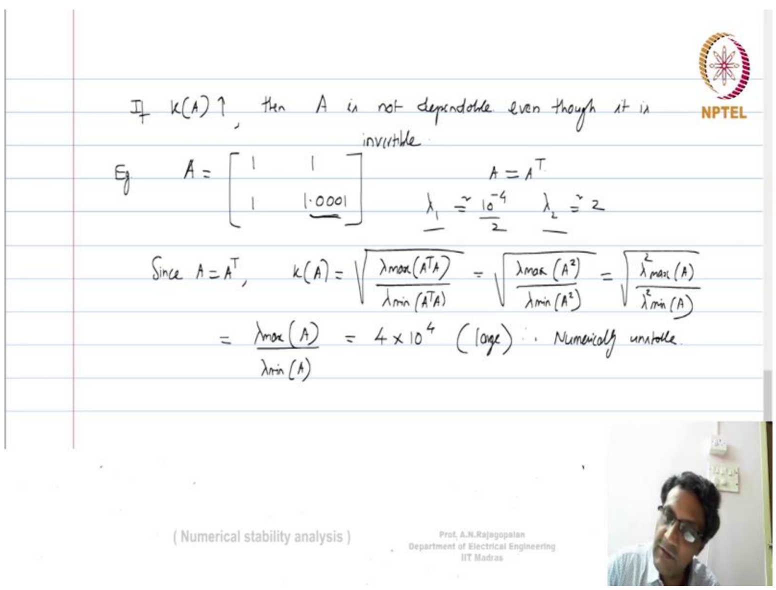
Therefore multiplying 1 and 2 we get norm delta x by norm x because the other one had like 1 by norm x here is less than or equal to so the sign does not change because all are norms here so therefore this sign would not change, norm A into norm A inverse into norm delta y then there you had actually 1 by norm y, now if you notice this is a relative change in y and this is the relative change in x this is in solution, let the change in solution, a little change in y or a little change in change in observation.

So, what we find is that the relative change in the solution is actually upper bounded by a number times the relative change in observation. Now, it is interesting that that norm A into norm A inverse we know is greater than or equal to norm A into A inverse which is equal to norm of i and the spectral norm of i, now at this point of time let us understand that these are all spectral norms, this is the spectral norms, it is are all two norms of these vector therefore, this is equal to and because the Eigen values of the of the identity matrix are all one therefore you know its singular values also one therefore norm of i is simply one.

Therefore this number it is always this product of this two norm, norm A is always a number that is greater than or equal to one what this amounts to saying is that that this relative change is actually upper bounded by this number times this therefore this number acts like an amplification factor, the larger this number is the higher will be the change in the solution for a change in the observation, this is an amplification factor and this number is called the condition number, it is called the condition number of A, called the condition number of A or denoted as kappa of A equal to norm A into norm A inverse which will then be equal to sigma max of A and norm of A inverse is what?

Norm of A inverse is sigma max of A inverse, this is norm of A of sigma max of A and norm of A inverse is sigma max of A inverse that is nothing but 1 by sigma min of A, therefore this is equal to the ratio of the maximum and minimum singular values of A and therefore this is called, so this is the condition number and this is what will tell you the health of the matrix. So, even if your matrix is invertible but then if the spread of the singular values is very large, so this is like indicative of the spread of singular values.

(Refer Slide Time: 20:40)



Therefore the spread is high, so then it means that so if kappa is high, kappa is high then A is numerically unstable, it is not a dependable matrix, it is not really dependable matrix, that means when you try to take inverse and you want to accept the solution you may be in trouble even though, even though it is invertible, even though it may be invertible, even though it is invertible, if it is not is invertible then of course, your condition number will be simply infinity even though this is invertible.

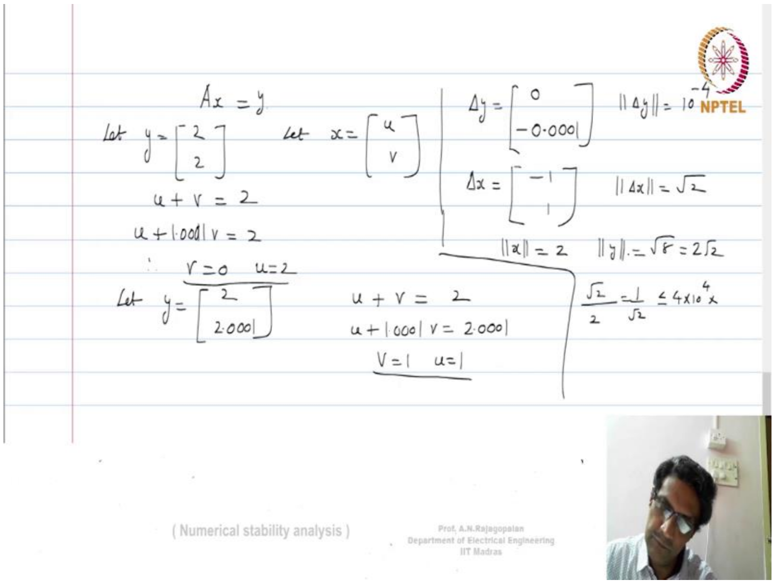
So, invertibility by itself is not enough that is the reason why I had a certain condition was kind of very very relevant, that talked about the continuity of solution with respect to observation and that is exactly what is captured by the condition number. Let me just take one example with regards to this and let us take this example that will tell you what, that can help you understand this, let us take a matrix A which is 1, 1, 1, 1.0001.

So, it looks like a contrived matrix because it was all 1, 1, 1, 1 then we know that we cannot even invert this guy but then a small number has been added to the last just to make it invertible in the hope that now that we have an invertible matrix all should be fine, we should be able to use this matrix but then as you would expect that would that would be a wrong move because this A as we can obviously see, if this was 1 then we could not even have inverted it and all that we have done is this last number is just a little bit away from 1 and therefore we expect that this A cannot be a numerically stable matrix.

Now, let us check this out, since A is equal to A transpose, so since A equal to A transpose, now first for A let us just look at fact that it is A is equals to transpose and its Eigen values if you look at it lambda 1, lambda 1 is approximately 10 power minus 4 by 2, lambda 2 is approximately 2 and since A is equal to A transpose you get the kappa of A if you notice if you could have remembered this route at lambda max of A transpose A by lambda min of A transpose A.

Now in this case because of the fact that A is equal to A transpose this becomes lambda max of A square my lambda min of A square, that means that is equal to root of lambda square max of A by lambda square min of A is equal to lambda max of A by lambda min of A, this happened because it was a special case of the symmetric matrix, this is equal to 4 into 10 power 4, if you divide lambda 1 by lambda 2, lambda 2 by lambda 1. So, the maximum Eigen value is 2 divided by minimum, so it is 4 into 10 power 4 this is very large therefore, numerically unstable matrix. Now what does that mean?

(Refer Slide Time: 24:12)

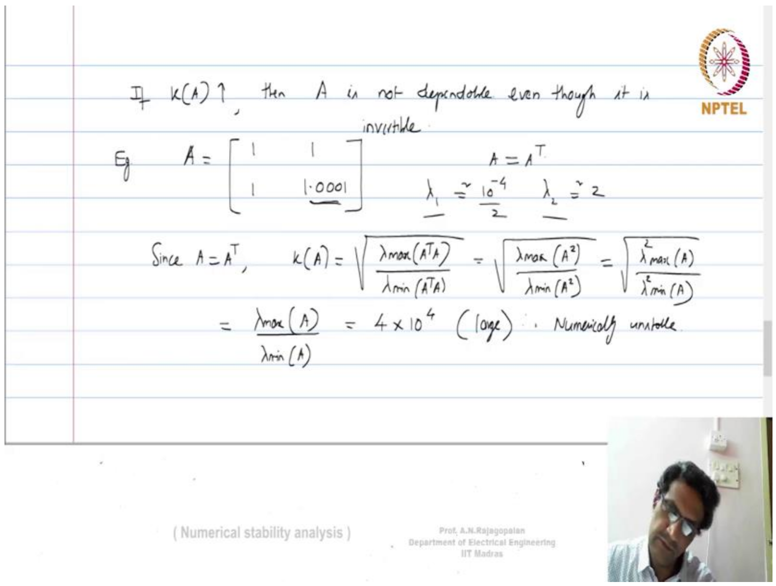


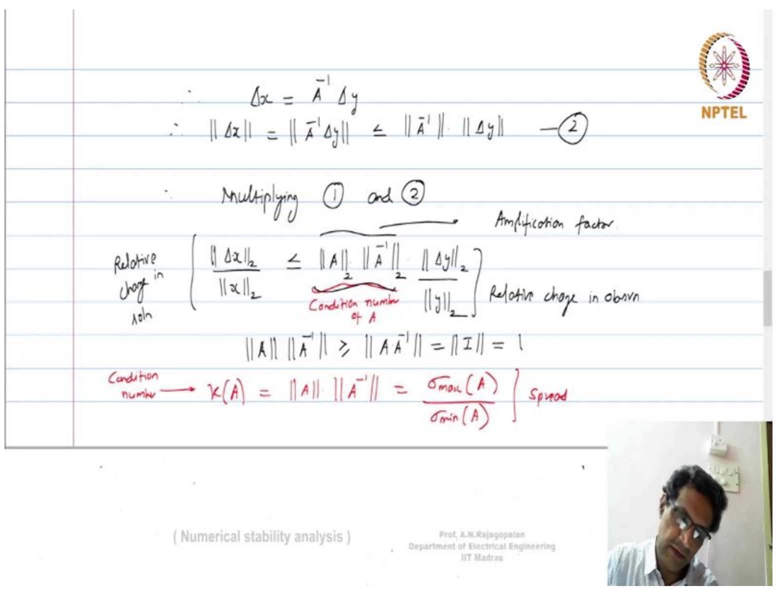
So, if you try to use this for some calculation it will go terribly wrong, say we want to solve Ax equal to b or Ax is equal to y. Now let for the same A, let y be equal to suppose somebody gives you y to be equal to let us say 2 and let x be equal to some u, v and if you solve for this you get u plus v is equal to 2 and u plus 1.0001 v is equal to 2 therefore it follows that v is equal to 0 and u is equal to 2.

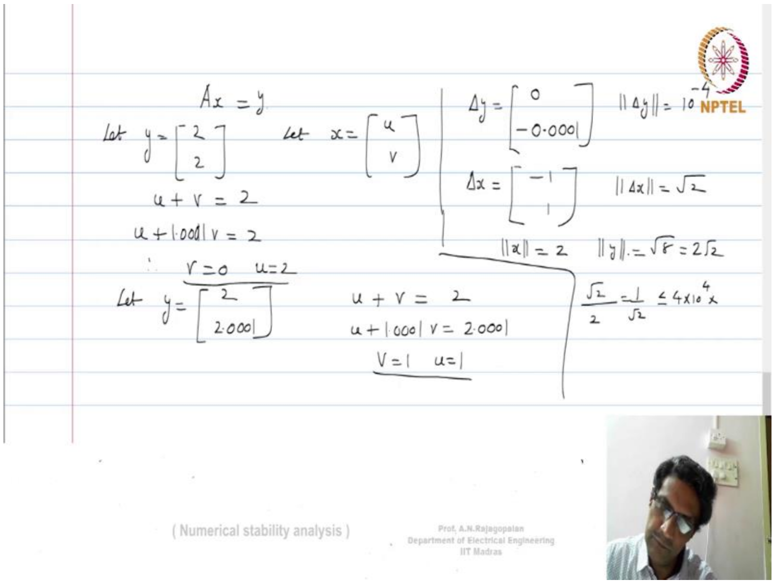
So, now let us say that my y changes a little bit to 2 and the 2.0001, now as you can see that y has hardly changed so y has just become, y has been changed by a very very small amount, now we would, I did not expect x to not change much but then then what will happen is x will indeed change and how does x change? So, if you do Ax equal to y for this you get and again u plus v is equal to 2 but now you will get u plus 1.0001 v is equal to 2.0001 or in other words you get v is equal to 1 and u is equal to 1.

So, see the change in u and v, earlier it was v equal to 0, u is equal to 2, now it is v is equals to 1, u is equal to 1 so which is a big change and what kind of relative change have we got here and if you want to check that out what could relative change. So, if you look at norm, so if you look at delta x now what is the change in, let us first look at delta y what was the change in delta y? So, delta y changed as 0 and minus 0.0001, what is the change in delta x? Delta x changed as minus 1, 1 because 0 became 1, 2 became, 0 became 1 and 2 became 1, therefore norm delta y is equal to 10 power minus 4, norm delta x is equal to root 2, what was norm x? Norm x was 2, so the norm x equal to 2, what was norm y? Norm y us root of 4 plus 4, 8 that is equal to 2 root 2.

(Refer Slide Time: 26:55)







Therefore according to our expression that we had here which was this in inequality it is so norm delta x by norm x if you see so norm delta x is root 2 by norm x which is 2 is equal to 1 by root 2 should be less than or equal to the amplification factor kappa, kappa of A was 4 into 10 power 4 into norm delta y by norm y, norm delta y by norm y.

So, norm delta y is 10 power minus 4 by norm y is 2 root 2, therefore this number is equal to root 2, whereas the one on the left is 1 by root 2, therefore clearly 1 by root 2 is less than or equal to root 2, so that satisfy the inequality but then what you also notice is that the role of (())(27:49) the role of the condition number is such that if kappa is small like kappa becomes low of course, you can definitely a minimum value 1 but as kappa increases then, if kappa decreases it automatically makes sure that this relative change will always be less than or equal to so the upper bound will automatically reduce.

Therefore that is the reason why you are having a kappa that is close to 1 is most ideal and do not think that the identity matrix for example here if we had an identity matrix that had a condition number, clearly the identity matrix has a condition number that is 1 but then there are in fact other matrices that also have a condition number that is 1, for example, unitary guys, if you take a unitary matrix because of the fact that u transpose u is equal to identity therefore it has a kappa, kappa u is in fact 1.

So, therefore it is not true that only the identity matrix will have a condition number is 1 wherein some other matrices can also have 1 as condition number and that would be the most ideal matrix to deal with but then even a higher number of condition number is acceptable provided it is not way too large, if it becomes way too large then it means that your solution becomes very sensitive to your observation, to any change in observation, therefore even if you have a matrix that is invertible will not help you because your condition number will be really high.