Computational Fluid Dynamics

Prof. Dr. Suman Chakraborty
Department of Mechanical Engineering
Indian Institute of Technology, Kharagpur

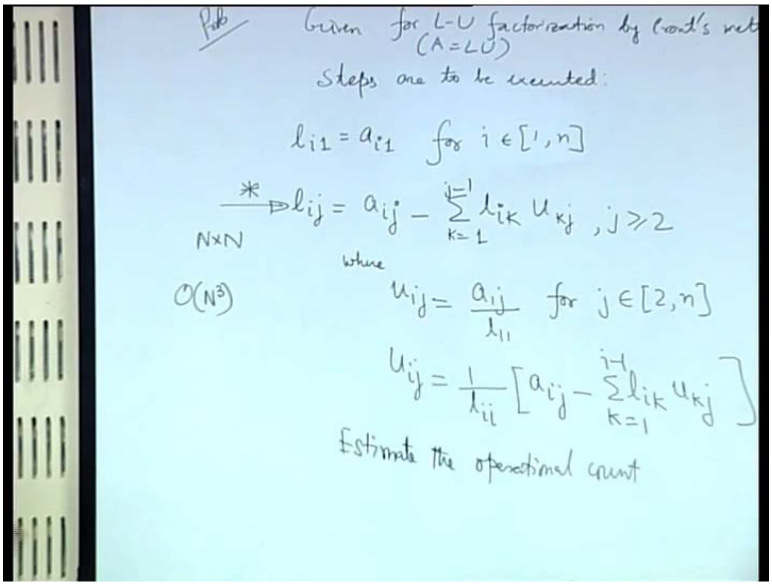
Module No. # 01
Lecture No. # 24

Elimination Methods: Error Analysis

So far, we have discussed some of the elimination methods for solving systems of linear algebraic equations. So, what are the methods that we have discussed? We have discussed the Gaussian elimination, LU decomposition method and the Tri-diagonal matrix algorithm. These are some of the main methods that we have discussed.

Now, what we will be interested to understand subsequently is that, once we use these methods, what are the important features of these methods that we must keep in mind, that is, what is the error that is possible to be incurred? As you are using a particular method, we will go into that, but before getting into that, let us also recall that there are several important features of a method. One is, of course, the error that we will be looking into subsequently, but the other is the efficiency of the method in terms of the operational counts. So, we have seen or we have assessed the efficiency of the methods, in terms of the number of operations for Gaussian elimination as well as LU decomposition. Now, we will work out one example, where we do that for the LU decomposition method before proceeding further.

(Refer Slide Time: 01:47)



So, given for LU factorization by Crout's method following steps are to be executed. First step is, l i1 is equal to a i1 for i equal to 1 to n. Second step, l ij is equal to a ij minus summation of l ik u kj for k equal to 1 2 j minus 1 where u ij is equal to a ij by 1 l1 for j equal to 2 to n and u ij is equal to 1 by l ii for j is equal to i plus 1 to n and j greater than equal to 2.

I have not written the steps in the order in which they need to be executed, that is what I leave on you as an exercise, but what we are interested from this is, estimate the operational count. You can see here that the calculation of l ij and u ij, they are mixed like for some l ij's. You calculate them earlier. Then you calculate some u ij, then some l ij like that. So, some components of l and some components of U. Basically, you are calculating simultaneously. Now, what is the order in which you calculate, that is, what I leave on you as an exercise, but what we intend to find out from here is, that what the operational count is. So, if a, has size of n by n, then what is the operational count of this 1, so that can be very easily assessed by looking into this system of equations. How do you assess that?

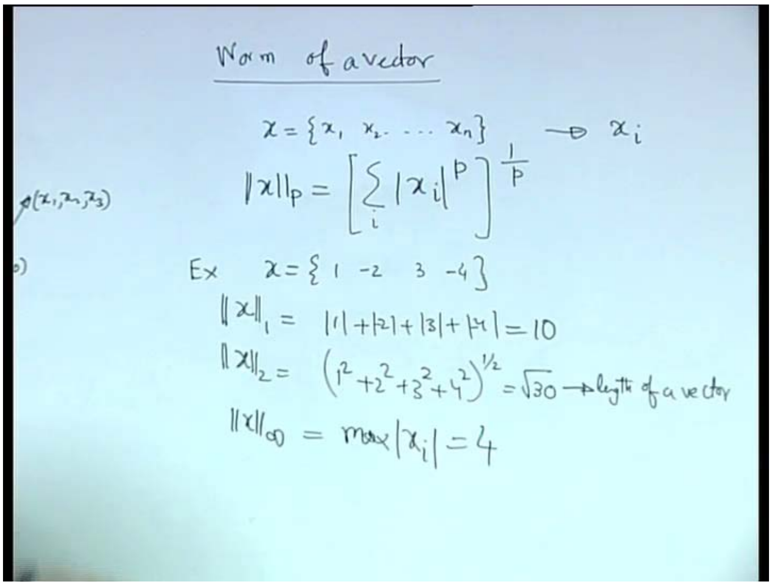
Here, you see that what are the indices over which the for loop is there. So, first there is an index i. So, look into this key step, l ij. This step will tell you that what the operational complexity is. You have i equal to 1 to n. Then, you have j, that is, of the order of n. It runs and k from 1 to j minus 1. So, that is also of the order of n.

So, you have one loop over i 1, loop over j 1 loop over k, each running upto of the order of n. So, that makes this particular step as the rate determining step or the number of count determining step, that is, of the order of n q. So, the whole idea is to get a feel that by looking into an algorithm, how can you figure out or even not a full algorithm just a skeleton of an algorithm. This is not a well described full algorithm. This is the skeleton of the algorithm which is good enough to tell us that, what would be the number of operations.

Next, we will move into the main agenda of today's lecture, that is, error analysis for the elimination methods. So, what we intend to analyze? We know that every numerical method is susceptible to certain errors. We have earlier seen some typical errors like truncation error, where you have the error due to truncation of the Taylor series upto a finite number of terms. Then, round of error because of rounding of the number upto a particular number of decimal points.

So, there are certain errors which are intrinsic to the method itself and there are certain errors which are intrinsic to the computing platform that you are using. Your numerical error is a combined consequence of these two. Now, when we talk about the error in any of the elimination methods, we will look into the errors, which are intrinsic to the method. To do that, we will take up certain examples, but we will introduce an important concept or important terminology before proceeding with the error analysis and that is, the norm of a vector or a matrix.

(Refer Slide Time: 09:19)



First, we will consider norm of a vector. Let us say that, you have a vector x with elements as x 1, x 2 upto x n. It may be a rho vector like this or it may be a column vector. So, symbolically an element of x is given by x i, where i varies from 1 to n. The pth norm of the vector x is given as follows. First, you take the mod of x i to the power p, sum it up over all i, then raise the whole thing to the power 1 by p.

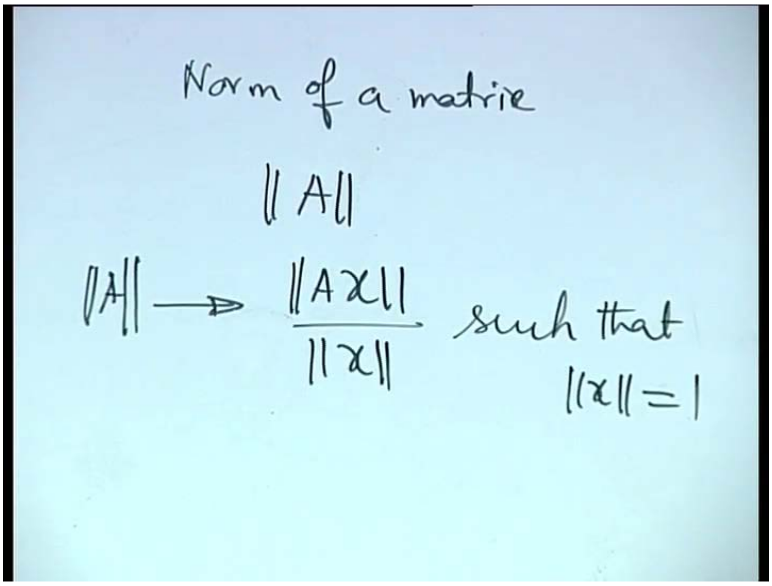
Let us take an example to illustrate this. Say, x is equal to 1 2 3 4. This vector or let us put some minus signs 1 minus 2 3 and minus 4, the first norm summation of mod of x i to the power p. So, mod of 1 plus mod of 2 plus mod of 3 plus mod of 4 is 10. Second norm, it is 1 square plus 2 square plus 3 square plus 4 square to the power of half, whatever it is, 16 plus 925 plus 429 plus 130. So, square root of 30.

So, this is called as length of a vector. You can see that physically in a 3 dimensional space, if the vector had only 3 entries, then square root of x 1 square plus x 2 square plus x 3 square would have represented the length of the vector. So, for example, if you had origin 0 0 0 and a point with coordinates x 1, x 2, x 3, then the length of this vector is given by square root of x 1 square plus x 2 square plus x 3 square.

Here, in the linear algebra, we are not always dealing with a physical vector. We are dealing with a vector in an algebraic sense. So, it is a bit more abstract than a physical vector, but it is not abstract in a true sense. In a way, that it is a straight forward extension of the physical understanding of the geometrical vectors in n dimensional spaces, where you really cannot geometrically conceive the vector. You can use a similar understanding, if you reduce the dimensionality into 2 dimensions or 3 dimensions.

Now, in this way, you can calculate the norms. Let us look into the infinity norm. Now, for infinity norm, there is a special definition. What is the definition? Here, when you have each element of the vector, you raise it to the power of infinity. So, out of this, you have many terms. Which term will dominate out of these? Whatever is the largest in magnitude, infinity power of that, will dominate over all other terms. So, effectively it will become the largest number magnitude to the power infinity, that to the power 1 by infinity. So, that means, it is nothing, but you find out, what is the maximum magnitude of x i and that is, the infinity norm. So, that is 4. There are different other possible definitions of norms. We will not go into that. We will consider this as one of the very basic definitions.

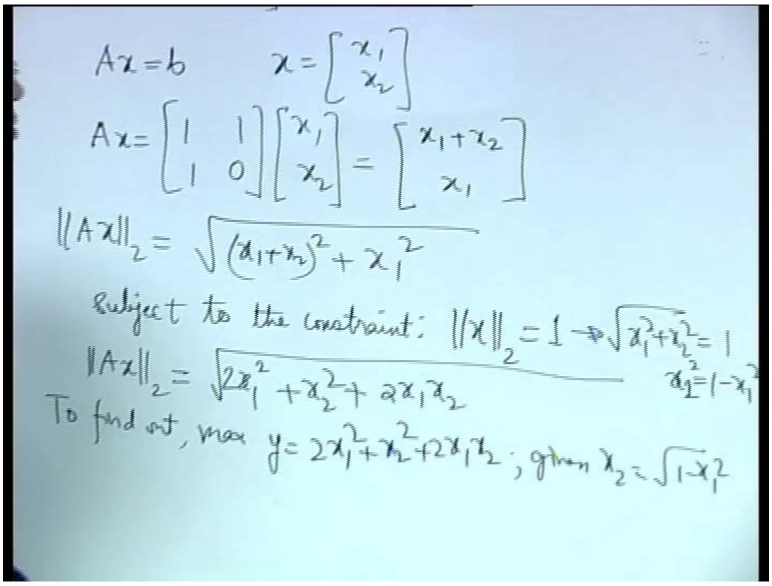
(Refer Slide Time: 15:28)



We will, next try to extend is to the concept of norm of a matrix. You can clearly see that this definition of norm that we have seen for a vector, it cannot be directly extrapolated to that of a matrix because matrix is not a linear structure or to be more precise, not a 1 dimensional structure like a vector.

So, usually, when you are interested to find out the norm of a matrix, it is defined through the norm of a vector. So, how it is defined? You introduce a vector x, such that you find out norm of x. So, you may define norm of A, as norm of Ax by norm of x. May be maximum of that or you may define it in a way such that norm of x is equal to 1, that is you may normalize it in this way. So, there are several possible definitions which are there. This is one of the acceptable definitions that you consider. Any arbitrary vector x, you find out norm of Ax. You divide that by norm of x, such that norm of x equal to 1. Then, the corresponding norm of Ax will be the Norm of A.

(Refer Slide Time: 18:02)

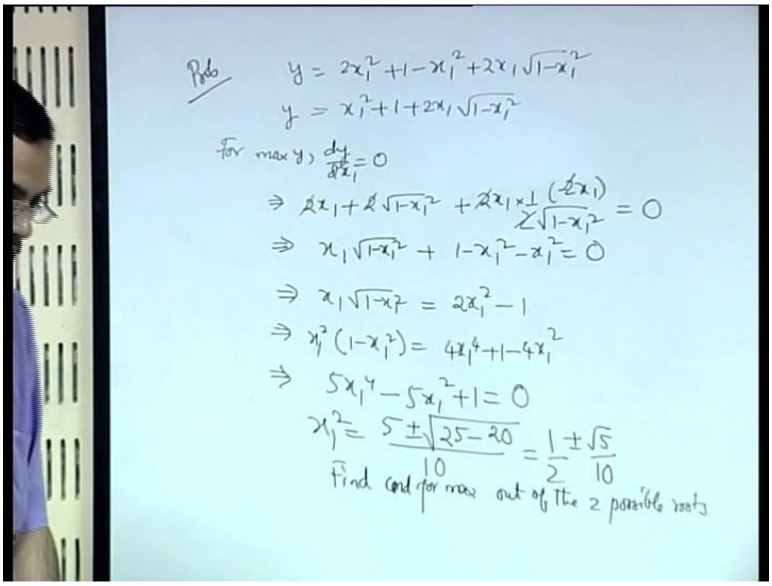


Let us take an example to illustrate this. So, this A is given. You have to find out the second norm of A. The first norm and the infinity norm for a matrix, these are kept with special purposes. We will touch upon them separately, but here, we consider the second norm. So, let us say that we have Ax equal to sum b, where x equal to x 1 x 2. So, what is Ax? x 1 plus x 2 and x 1. So, the second norm of Ax will be square root of x 1 plus x 2 whole square plus x 1 square subject to the constraint that mod of x, sorry norm of x 2 equal to second norm of x is equal to 1. So, square root of x 1 square plus x 2 square is equal to 1.

Now, if you find this one, you can eliminate x 2 and write it in terms of x 1 or you can eliminate x 1, write it in terms of x 2, but you will see that, it is non-ambiguous. Let us go through that and we will easily find out. So, norm of x 2 is equal to x 1 square. Here, this is 1 x 1 square another x 1 square, so 2x 1 square plus x 2 square plus 2x 1 x 2.

Now, you can eliminate x 2 and write x 1 as x 1 square as 1 minus x 2 square or x 2 square as 1 minus x 1 square. Still, x 1 will be a variable. So, it will be non-ambiguous depending on the choice of x 1. There will be different possible values to resolve this ambiguity. What we do is, here we consider max of this one in the definition. So, norm of A, we consider max of norm of Ax by norm of x, such that norm of x equal to 1. So, effectively we have to find out max of say, y equal to 2x 1 square plus x 2 square plus 2x 1 x 2. Given x 2 is square root of 1 minus x 1 square.

(Refer Slide Time: 23:26)



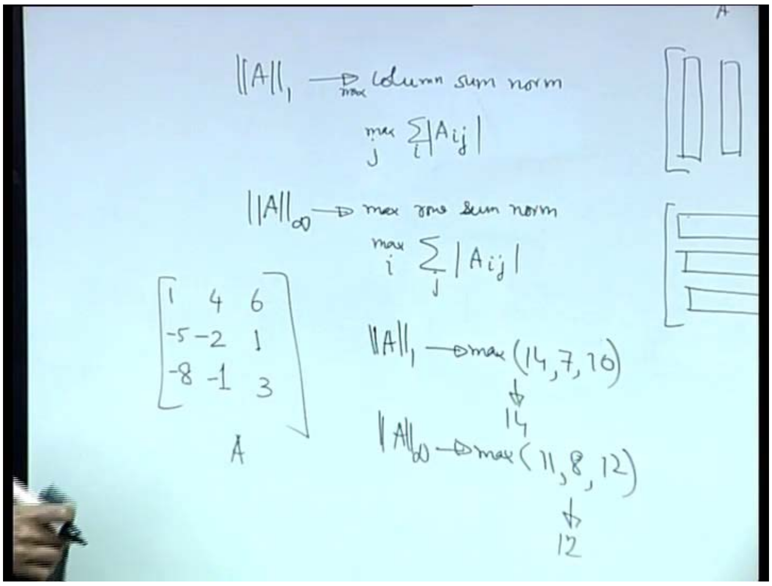
So, y is 2x 1 square plus x 2 square is 1 minus x 1 square plus 2x 1 into square root of 1 minus x 1 square. So, that is x 1 square plus 1 plus 2x 1 square root of 1 minus x 1 square. So, for maximum y dy dx 1 must be equal to 0. That means 2x 1 plus 2 square root of 1 minus x 1 square plus 2x 1 into 1 by 2 root 1 minus x 1 square into minus 2x 1 equal to 0.

So, that means, x 1 square root of 1 minus x 1 square plus 1 minus x 1 square minus x 1 square is equal to 0. So, it means, x 1 square root of 1 minus x 1 square is equal to 2x 1 square minus 1. You can take square of both sides. So, it becomes x 1 square into 1 minus x 1 square is equal to 4x 1 4 plus 1 minus 4x 1 square. So, it is 5x 1 4. Then, this is minus 4x 1 square and minus x 1 square. So, 5x 1 square plus 1 is equal to 0. So, x 1 square is equal to, it is a quadratic equation in x 1 square. So, you can solve for x 1 square. So, x 1 square is 5 plus minus 25 minus 20 by 10.

Now, can we take both the signs? Both are positive, but just check, whether we have made any calculation mistake or not. So, this should be fine. Now, you also have to check, whether this corresponds to a maxima or a minimal. There are 2 possible solutions for x 1 square. Out of these, it is not trivial that both will be maximum. So, we have to find out the corresponding. I am leaving that on you as an exercise. So, find condition for maxima out of the 2 possible roots. That will give you the corresponding value of x 1 square and from that, you can calculate x 1. Hence, x 2 and hence the elements of a x.

So, effectively what you have to calculate? You have to calculate max of the norm of second norm of A x. Given that norm of x equal to 1, second norm of x equal to 1. So, we have done the majority of the steps. It is now just to find out what is the corresponding condition for the maximum and substituting that in the second norm A x. Now, as I mentioned that there are certain norms which are preserved for special meaning for matrices. Those are the first norms and the infinity norms.

(Refer Slide Time: 30:27)



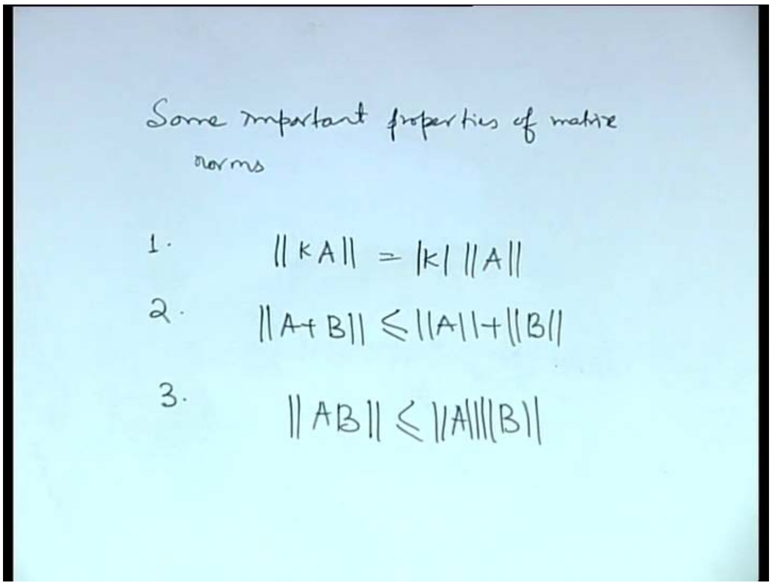
So, the first norm of A is called as the column sum norm. So, what you do by this? If you have a matrix A, you find out the norms of the vectors. Which norm you find out? We will find the infinity norms of the vectors which are the columns of this particular matrix. So, essentially what you do? Essentially, you find out some of the magnitude of the elements of each column and find out of that which one is the maximum. So, this we can call as maximum column sum norm.

So, summation of A ij, summation over i. You have to find out maximum for which j? It is there. Of course, with a mod, we will consider some numerical examples to illustrate this one, but let us first go to the definition. Similarly, the infinity norm is called as maximum rho sum norm. So, here, what you do? You consider the rows rho vectors which constitute the matrix for each rho vector. You find out again, sum of the magnitude of A ij. So, you find the sum over j and find out for which rho, it is maximum. Then, that maximum number is corresponding rho sum norm.

So, let us consider a numerical example. Let us say, we have the entries of the matrix like, let us say we have a matrix like this. So, when we consider, let us say, this is matrix A. When we consider the first norm of A, then what are the options that we have from the column 1. We have 1 plus 5 plus 8. So, 14 from the column 2, 4 plus 2 plus 1 from the column 3, 6 plus 1 plus 3, so it is 14 the infinity norm from the row 1, we have 6 plus 4 plus 1. So, 11 from row 2 1 plus 2 3 plus 5 8 from row 3 3 plus 1 4 plus 8 12, so it is 12.

These are some of the limiting norms for matrices very easy to calculate. You have just witnessed that, it is a bit tedious to calculate the norms, other than the first norm and the infinity norm for the matrix. Even, for a simple 2 by 2 matrix, you found out that we had to go through a reasonable number of steps to find out the second norm of a matrix. Whereas, the first norm and the infinity norm can be calculated in a much more simple way and because of that and because of the fact that, they give some limiting values of the norms. These norms are very commonly used for error analysis in numerical methods.

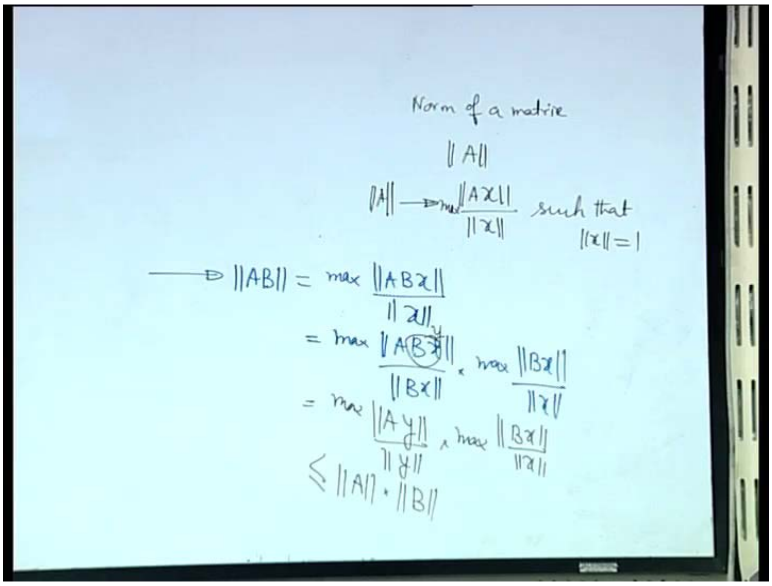
(Refer Slide Time: 36:39)



Now, next we will try to learn some important properties of these norms of matrix norms, one of the very interesting properties is that norm of kA, where k is A scalar. How it is related to norm of A? Is it k or you require A mod of that? See, you have seen through the example of the norm calculations that we have come across. That norm essentially, ends up with something positive. The second observation is that, it is something more general than mod. So, the mod is like the second norm, but you could have different types of norms possible, but like mod, it returns always a positive number like k is A scalar. It is a sign scalar. It can be positive negative, whatever. So, this mod of k is there to preserve the positivity of norm of kA.

Then, norm of A plus B is less than equal to norm of A plus norm of B. This is also a very important property. There is a third property, which we will be using more often for error analysis, that is norm of AB is less than equal to norm of A into norm of B.

(Refer Slide Time: 39:21)



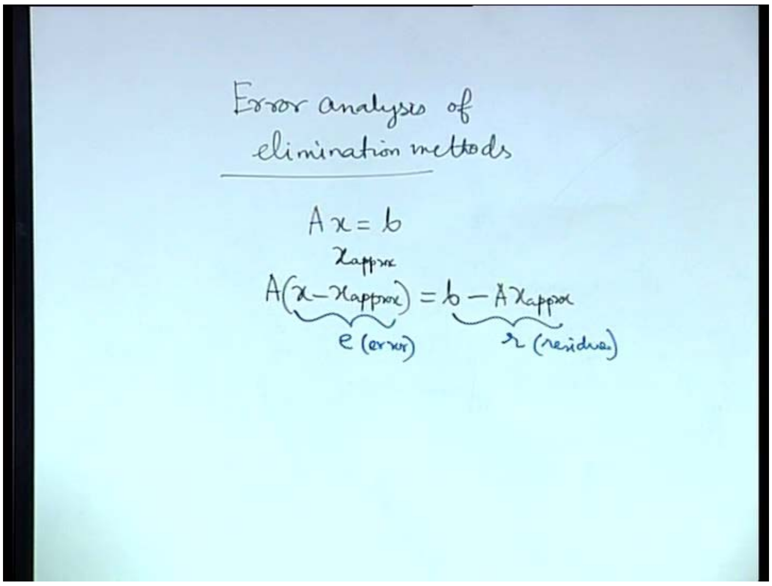
So, this can be shown in a simple way. You can write norm of AB as max of norm of ABx by norm of x. This you can also write as max of norm of ABx by norm of Bx into max of norm of Bx by norm of x as if you have multiplied and divided by norm of Bx.

You can consider Bx as A vector y. So, this is max of norm of Ay by Norm of y into max of norm of Bx by norm of x. Remember, that there is no sanctity in the names of x y like that. These are just some arbitrary vectors. So, this must be less than equal to norm of A into norm of B. The first part belongs to norm of A and the second part belongs to norm of B.

So, you can use this inequality very effectively for calculation of errors in elimination methods. Let us now go into that. You can see that, it is, you have not put any constraint on the norm of x or norm of y. So, whatever it is, whatever is norm of A, it is always less than the maximum of this possible value because out of all the possible maximum values, you could always have the norm of A. It cannot go beyond the maximum possible value of this. It is not just the maximum possible value of this; it is with a constraint that norm of y equal to 1. So, out of all the possible maximum values, it cannot supersede the maximum possible values. So, it could have several of the maximum possible values, with different choices of norm of y or norm of x.

So, you can always normalize norm of x or norm of y in a certain way as per your convenience. Whatever is the manner in which you normalize, you always have the value less than the maximum because maximum is the maximum you cannot exceed that one.

(Refer Slide Time: 43:10)

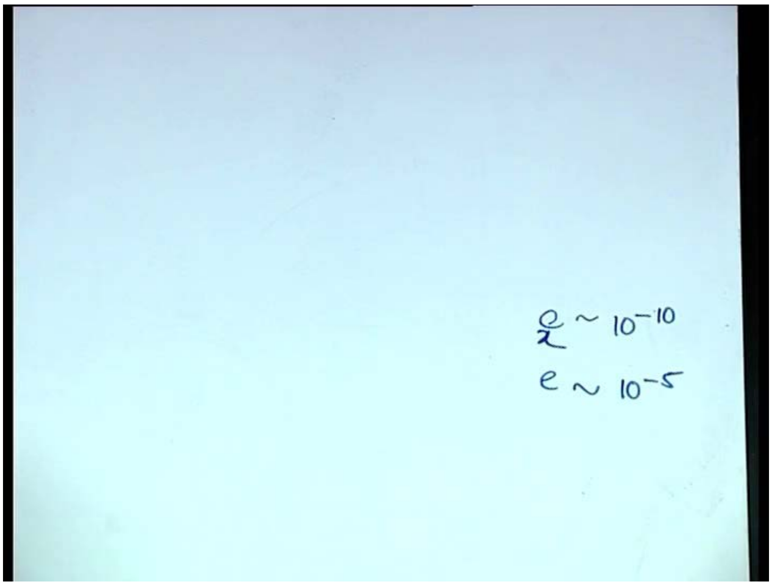


Next, we try to utilize these concepts for the error analysis of elimination methods. So, let us say, that we have a system of equations Ax equal to b, where A is the co-efficient matrix, x is the vector of unknowns and b is the right hand side. Let us say, that x approximate with A approximate numerical solution.

So, A into x minus x approximate is equal to b minus A into x approximate. What is x minus x approximate? That is the error in x right. So, we call that as e and b minus A approximate x A into x approximate. This is called as a residue r. If the approximate solution was same as the actual solution, then the residue would be 0. So, residue is also an indicator of the deviation of the actual solution from the approximate solution.

Now, the question is, what do we intend to do in the error analysis? In the error analysis, we want to highlight what is an estimate of an upper bound of the error, that is, what can be the maximum possible error incurred. Always remember, that why do we need an error analysis? You are coming up with an approximate solution. You do not know the exact solution because if you know the exact solution, why would you go for an approximate solution. So, obviously, you do not know the exact solution. You are going for an approximate solution without knowing the exact solution. You are interested to make an estimate of what? It is the deviation from the exact solution, that estimation, not the exact quantification, but an estimation of that is the objective of your error analysis. Now, it is not legitimate. If we just consider the error as a magnitude by itself. For example, let us say, that you are in one case. You are having error in x of the order of 10 to the power minus 10.

(Refer Slide Time: 46:56)

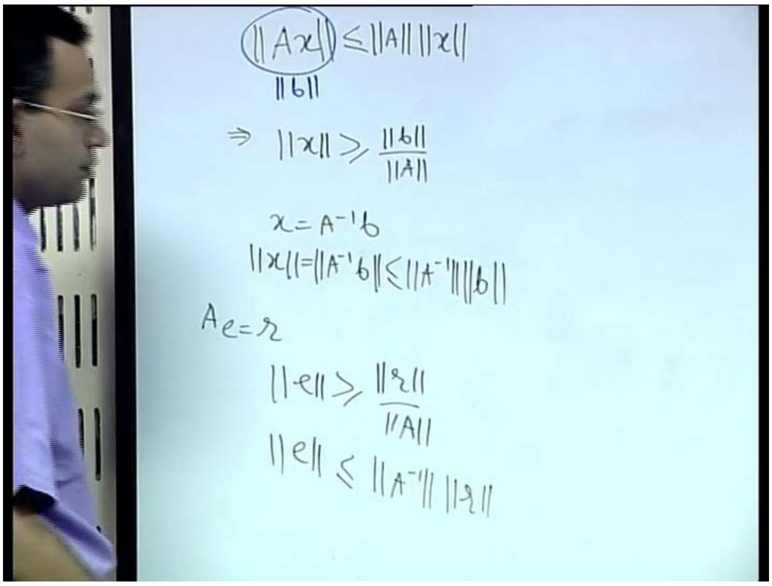


In one case, you have e of the order of 10 to the power minus 10. In another case, you are having e of the order of 10 to the power minus 5. Now, looking at these, can you tell in which case the error is more or in which case the error is less? You cannot really tell because it depends on, what is x as compared to which you are quantifying the error. It is the relative error, that is important and not the absolute error. Why?

For example, let us say, that you are dealing with a problem, where you are having dimensions of the order of atomic lengths scales. So, of the order of angstroms, then with respect to that 10 to the power minus 10 may not be a negligible one. Whereas, in another case, you are dealing with lengths of meters with respect to 10 to the power minus 5 with meter with respect to 1 meter, may be negligible.

So, what is your x? So, if x is of the order of 10 to the power minus 10, definitely e of the order of 10 to the power minus 10 is not negligible. If x is of the order of 1, then e of the order of 10 to the power minus 5 may still be negligible. So, what is negligible or what is not negligible, it depends on what is the magnitude of the number itself that you are considering. For which you are estimating the error, that is, the first thing that we have to keep in mind. So, it is not e as such which is important, it is the relative magnitude of e with respect to x. That is what is important.

(Refer Slide Time: 48:49)



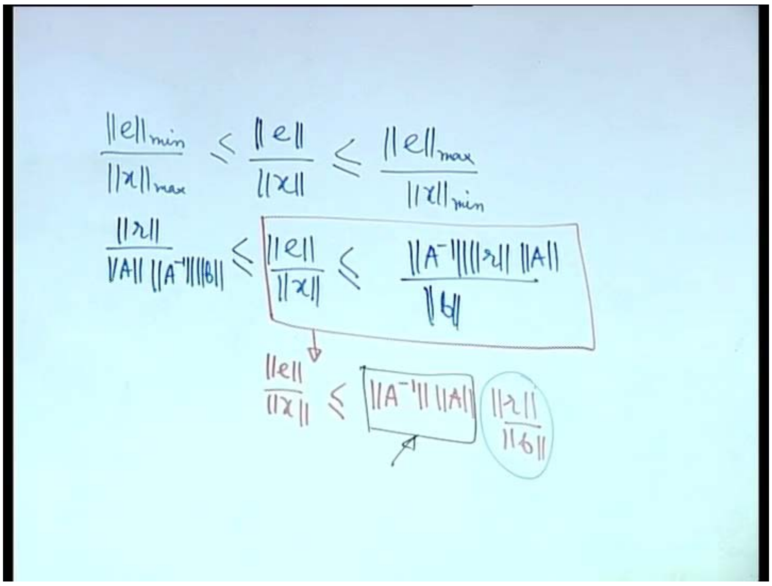
So, what is our point of interest or matter of interest? It is norm of e by norm of x. It is an indicator of the relative error. Norm is a more general quantity than the mod. So, we use the norm for estimating the error. So, we will find out, we will try to find out an upper bound of this one that will be the objective of our subsequent analysis.

To do that, we can recall that norm of x is less than equal to norm of A into norm of x. What we used for 2 vectors A and b is, also true for 1 vector A to matrices A and b. Whatever we used, the same is true for one matrix A and one vector x. So, norm of Ax less than equal to norm of A into norm of x and norm of Ax is equal to norm of b because Ax is equal to b.

So, from here, we can see that norm of x is greater than equal to norm of b by norm of A. Similarly, you can write x equal to A inverse b because Ax equal to b, you can write x equal to A inverse b. Therefore, norm of x equal to norm of A inverse b is less than equal to norm of A inverse into norm of b.

So, we get one lower bound of x and one upper bound of x, norm of x. Now, just you can write Ax equal to b. You can write A e equal to r, where r is the residue. So, this follows from Ax equal to b. So, if you have A e equal to r, then you can just replace x with e and b with r. So, you have norm of e greater than equal to norm of r by norm of A and norm of e is less than equal to norm of A inverse into norm of r.

(Refer Slide Time: 52:17)

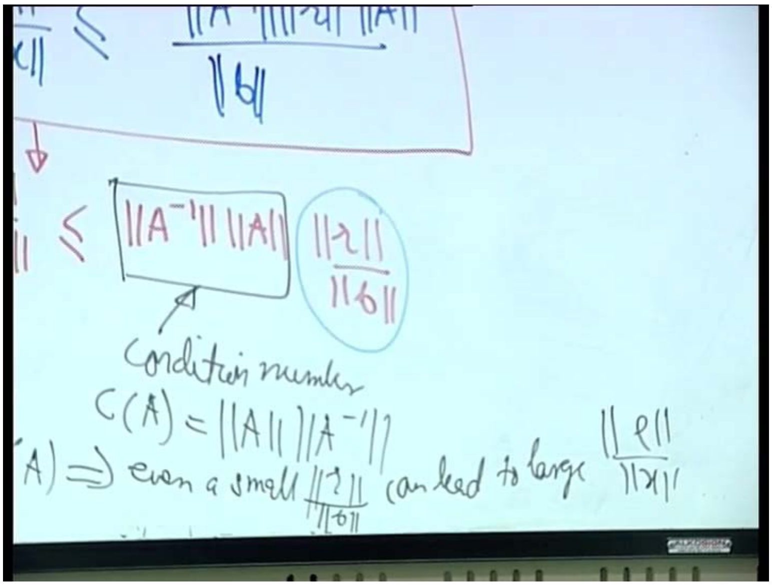


Now, let us go into the bounds of norm of e by norm of x. So, norm of e by norm of x, it is less than equal to max of norm of e by minimum of norm of x and greater than equal to minimum of norm of e by maximum of norm of x. Of course, this lower limit is not important error. Greater than something is not important error. Less than something that is important. What is the upper bound of the error, that is, for a conservative design of the algorithm? That is more important.

So, let us substitute these ones. What is norm of e minimum? No. Out of these, we have now these 4 inequalities which w can use. Norm of r by norm of A. What is norm of x maximum? Norm of A inverse into norm of b. Then, let us go to the other side of the inequality norm of e max is norm of A inverse into norm of r and divided by norm of x mean. So, what is norm of x minimum? That is, norm of b by norm of A. So, now, we will focus on this inequality. So, what we can see is, the relative error norm of e by norm of x is less than or equal to norm of A inverse into norm of A into norm of r by norm of b.

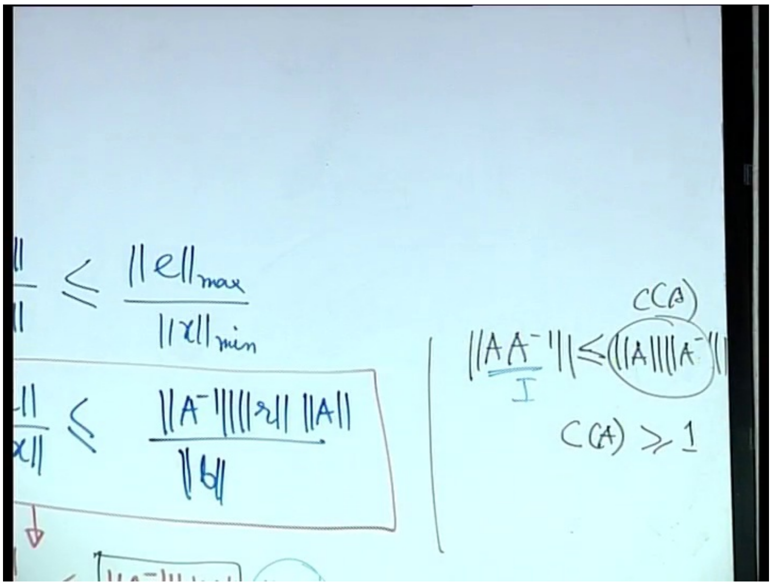
What we can conclude from here? We can conclude from here that, even if we have a small r by b, norm of r by norm of b r is the residual. So, if you have a small residual, many times we test the accuracy of a solution in terms of the residual because if the residual is 0, then it is accurate. So, if we have a residual with respect to b norm of the residual by norm of b, that is even if that is small, the relative error may be large. If norm of A inverse into norm of A is large. So, it is the largeness of norm of A into norm of A inverse. That is a very important consideration. Therefore, this determines the condition of the accuracy of the system of equations that you are solving.

(Refer Slide Time: 52:17)



So, this is called as a condition number C of A. This is norm of A into norm of A inverse. Large condition number means, even a small r by b, norm of r by norm of b can lead to large norm of e by norm of x. If such a case occurs, then it is called as an ill conditioned system. So, a large conditioned number is bad. It shows a great sensitivity of the error with the residue. If the residual is small, that is blown up because of the large norm of A into norm of A inverse. So, the co-efficient matrix should be such that, norm of A into norm of A inverse should be as small as possible. Now, how small it can be?

(Refer Slide Time: 57:36)



You have norm of A into A inverse less than equal to norm of A into norm of A inverse. Norm of, remember A into A inverse is identity matrix of the same order and norm of, that is 1. So, the condition number of the matrix, which is C A is greater than equal to 1, so closer to the 1. It is more and more away from the 1 is worse and worse in terms of magnifying the error.

So, the condition number of the matrix is a very critical parameter for estimating the error in the elimination method. We should have it as low as possible, so as to make sure, that the relative error is not magnified. We stop here today. Thank you.