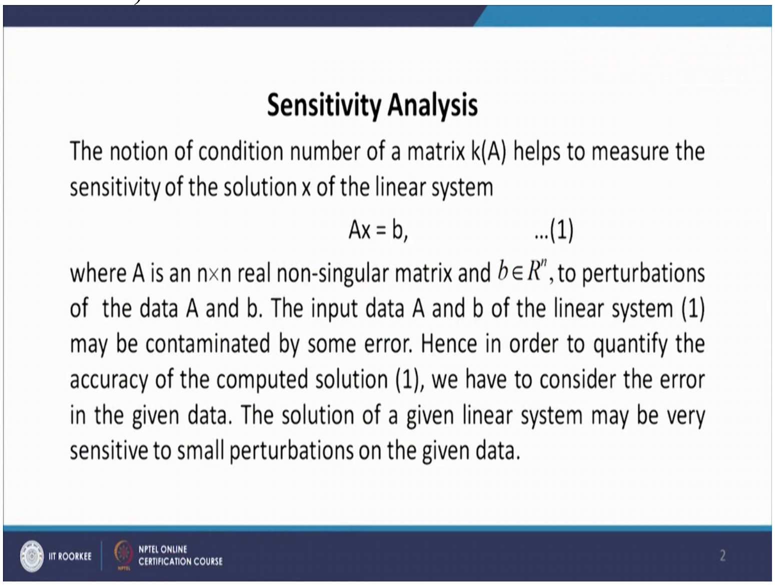
**Numerical Linear Algebra**
**Dr. P. N. Agrawal**
**Department of Mathematics**
**Indian Institute of Technology, Roorkee**

# Lecture - 37
# Sensitivity Analysis- I

Hello friends, I welcome you to my lecture on sensitivity analysis. There will be 2 lectures on this topic. This the first lecture, the notion of condition number of a matrix A, the condition number k, I mean.

(Refer Slide Time: 00:37)



## Sensitivity Analysis

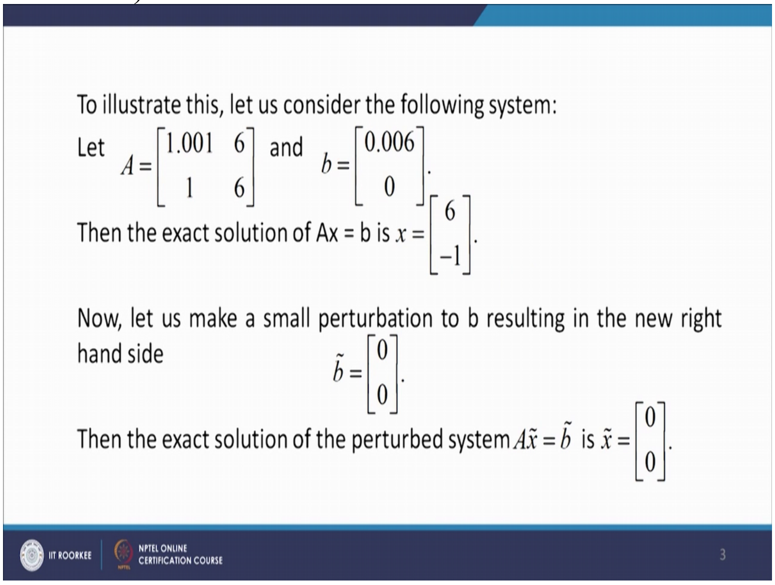
The notion of condition number of a matrix k(A) helps to measure the sensitivity of the solution x of the linear system

$$Ax = b, \qquad \ldots(1)$$

where A is an n×n real non-singular matrix and $b \in R^n$, to perturbations of the data A and b. The input data A and b of the linear system (1) may be contaminated by some error. Hence in order to quantify the accuracy of the computed solution (1), we have to consider the error in the given data. The solution of a given linear system may be very sensitive to small perturbations on the given data.

The notion of condition number of a matrix, say, k helps us to measure the sensitivity of the solution x of the linear system A x equal to b, where we are taking A to be n by n, a real non-singular matrix and b, a vector belonging to R to the power n to perturbation of the data A and b. The input A and b of the linear system may be contaminated by some error. Hence, in order to quantify the accuracy of the computer solution, we have to consider the error in the given data. The solution of a given linear system may be very sensitive to small perturbations on the given data.

(Refer Slide Time: 01:20)



To illustrate this, let us consider the following system:

Let $A = \begin{bmatrix} 1.001 & 6 \\ 1 & 6 \end{bmatrix}$ and $b = \begin{bmatrix} 0.006 \\ 0 \end{bmatrix}$.

Then the exact solution of Ax = b is $x = \begin{bmatrix} 6 \\ -1 \end{bmatrix}$.

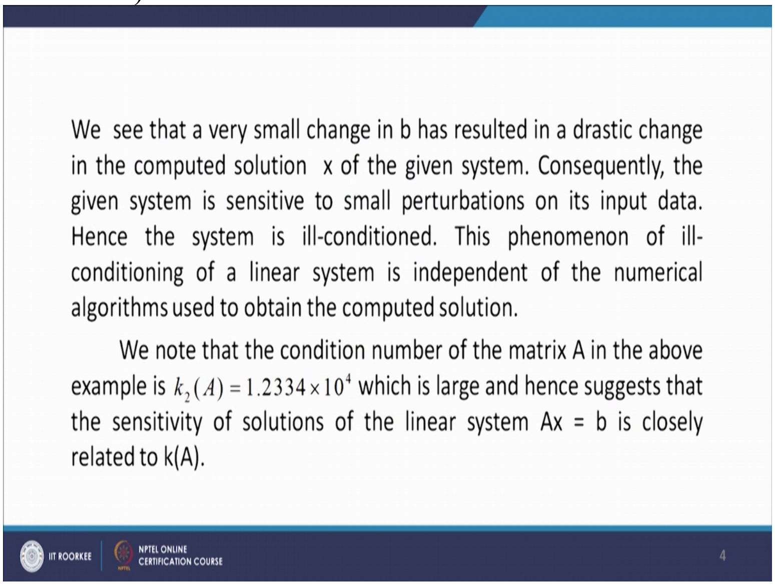
Now, let us make a small perturbation to b resulting in the new right hand side $\tilde{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Then the exact solution of the perturbed system $A\tilde{x} = \tilde{b}$ is $\tilde{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

To illustrate this, let us consider the following system. Let A, b is equal to the 2 by 2 matrix 1.001 6 1 6 and b is equal to the column matrix 0.006 and 0. Then, the exact solution of A x equal to b, as you can verify is, x equal to 6 minus 1.

Now, let us make a small perturbation to b, resulting in the new right-hand side, b equal to 0 0. So, when you take b equal to 0 0, you get, A x equal to A x cap equal to b x b cap. The exact solution of this perturbed system; A x cap b equal to b cap is x cap equal to 0 0.

(Refer Slide Time: 02:07)



We see that a very small change in b has resulted in a drastic change in the computed solution x of the given system. Consequently, the given system is sensitive to small perturbations on its input data. Hence the system is ill-conditioned. This phenomenon of ill-conditioning of a linear system is independent of the numerical algorithms used to obtain the computed solution.

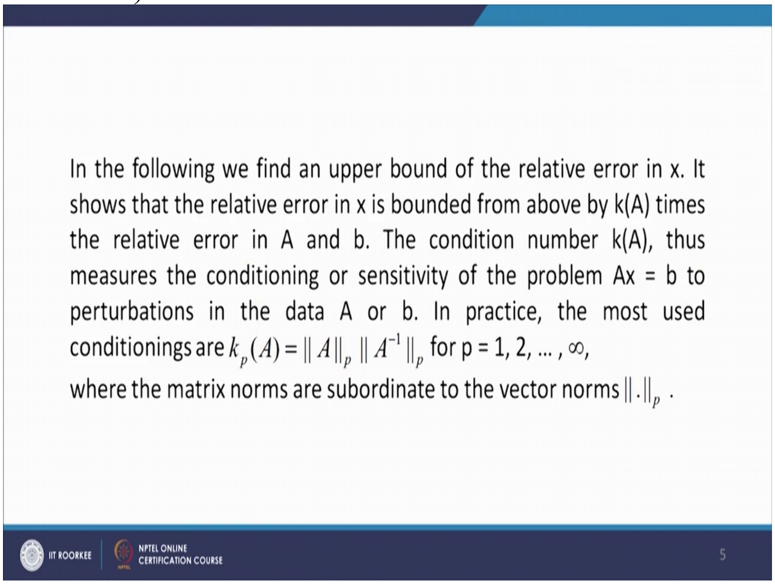
We note that the condition number of the matrix A in the above example is $k_2(A) = 1.2334 \times 10^4$ which is large and hence suggests that the sensitivity of solutions of the linear system Ax = b is closely related to k(A).

So, we can see that, when we make a very small change in b, a small change in b has resulted in a drastic change in the computed solution x of the given system. Earlier, the x was 6 minus 1. Now, the new x is 0 0; consequently, the given system is sensitive to a small perturbations on it is input data.

Hence, the system is ill conditioned. This phenomenon of ill conditioning of a linear system is independent of the numerical algorithms used to obtain the computed solution. Let us note that, the condition number of the matrix A in the above example is k 2 A equal to 1.2334 into 10 to the power 4, which is large. And hence, suggest that the sensitivity of solutions of the linear system A x equal to b is closely related to the condition number k A.

Now, we are going to discuss a theorem, where we shall find an upper amount of the relative error in x.

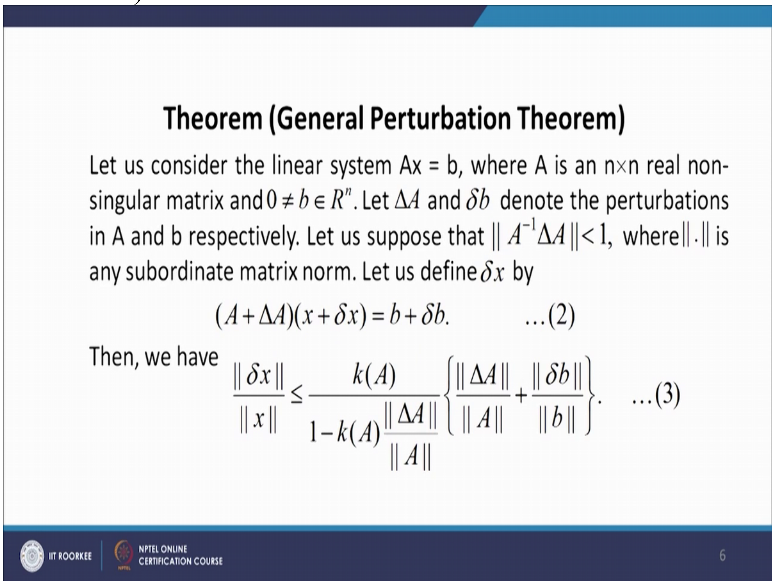
(Refer Slide Time: 03:17)



In the following we find an upper bound of the relative error in x. It shows that the relative error in x is bounded from above by k(A) times the relative error in A and b. The condition number k(A), thus measures the conditioning or sensitivity of the problem Ax = b to perturbations in the data A or b. In practice, the most used conditionings are $k_p(A) = \|A\|_p \, \|A^{-1}\|_p$ for p = 1, 2, … , ∞, where the matrix norms are subordinate to the vector norms $\|.\|_p$ .

It will show that, the relative error in x is bounded from above by the condition number of A, that is, k times the relative error in A and b. The condition number k A, thus measures the conditioning or sensitivity of the problem A x equal to b to perturbations in the data A or b in practice the most used conditioning are k p A equal to norm of A p into norm of A inverse p, where p is equal to 1 to n; so and infinity where the matrix norms are subordinate to the vector norms norm p.

(Refer Slide Time: 03:51)



**Theorem (General Perturbation Theorem)**

Let us consider the linear system Ax = b, where A is an n×n real non-singular matrix and $0 \neq b \in R^n$. Let $\Delta A$ and $\delta b$ denote the perturbations in A and b respectively. Let us suppose that $\| A^{-1}\Delta A \| < 1$, where $\| \cdot \|$ is any subordinate matrix norm. Let us define $\delta x$ by

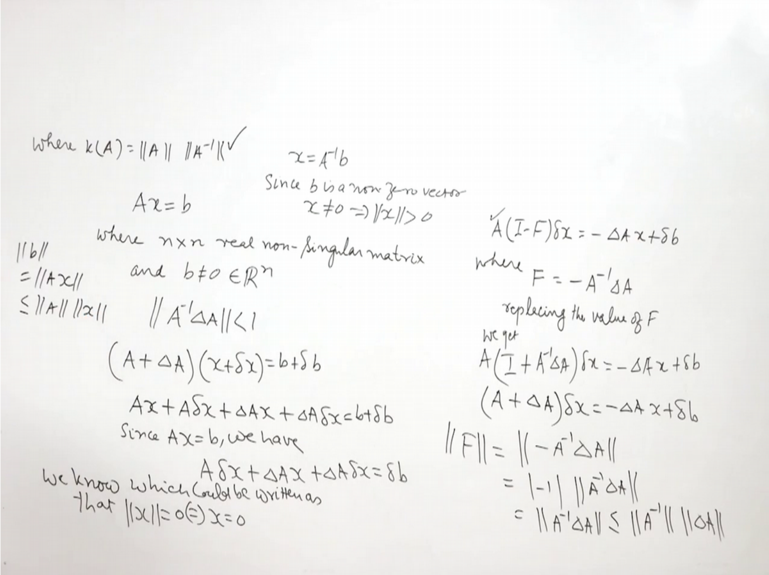
$$(A+\Delta A)(x+\delta x) = b+\delta b. \qquad \ldots(2)$$

Then, we have

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{k(A)}{1-k(A)\dfrac{\|\Delta A\|}{\|A\|}}\left\{\frac{\|\Delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right\}. \qquad \ldots(3)$$

Now, let us consider the linear system A x equal to b, where A is an n by n real non-singular matrix and 0 is not equal to b belonging to R n.

(Refer Slide Time: 04:00)



where $k(A) = \|A\| \, \|A^{-1}\|$ ✓

$x = A^{-1}b$

Since b is a non zero vector $x \neq 0 \Rightarrow \|x\| > 0$

$Ax = b$

where n×n real non-singular matrix and $b \neq 0 \in R^n$

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

$\|A^{-1}\Delta A\| < 1$

$(A+\Delta A)(x+\delta x) = b+\delta b$

$Ax + A\delta x + \Delta A x + \Delta A \delta x = b + \delta b$

Since $Ax = b$, we have

$A\delta x + \Delta A x + \Delta A \delta x = \delta b$

which could be written as

✓ $A(I-F)\delta x = -\Delta A x + \delta b$

where $F = -A^{-1}\Delta A$

replacing the value of F we get

$A(I + A^{-1}\Delta A)\delta x = -\Delta A x + \delta b$

$(A+\Delta A)\delta x = -\Delta A x + \delta b$

$\|F\| = \|-A^{-1}\Delta A\|$

$= |-1| \, \|A^{-1}\Delta A\|$

$= \|A^{-1}\Delta A\| \leq \|A^{-1}\| \, \|\Delta A\|$

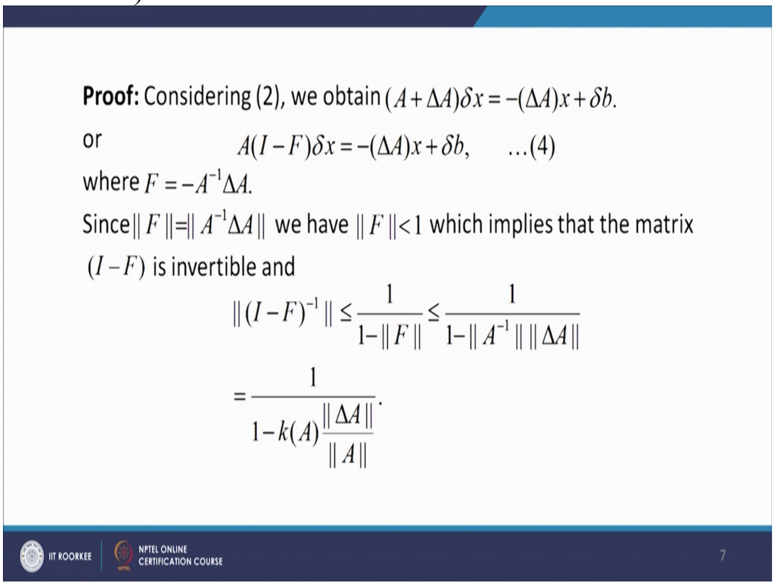
we know that $\|x\| = 0 \Leftrightarrow x = 0$

So A x equal to b, A is n by n real non-singular matrix and b is a non-0 vector in R to the power n. Now, let us consider delta A and delta b to be the perturbations in A and b respectively. Let us suppose, further that, norm of a inverse delta A is less than 1. The matrix norm of inverse delta is less than 1, where norm is the matrix norm, which is an subordinate matrix norm. As you know, the subordinate matrix norm means, norm of A b is less than or equal to norm of A into norm of b.

So, let us define delta x by A plus delta x A plus delta A into x plus delta x is equal to b plus delta b. So, then, we have norm of delta x over norm of x. The relative error in x is less than or equal to k, the condition number of a divided by 1 minus k, which is condition number A into norm of delta A over norm of A into the relative error in A which is norm of delta A over norm of A plus the relative error in b which is norm of delta b over norm of b.

So, here you can see that, the relative error in x depends on the condition number. Even if the relative errors in A and b are small, the condition never number plays a significant role. If it is very large, then the relative error in x could be also large. So, let us see how we prove the theorem.

(Refer Slide Time: 06:06)



**Proof:** Considering (2), we obtain $(A+\Delta A)\delta x = -(\Delta A)x+\delta b.$

or $$A(I-F)\delta x = -(\Delta A)x+\delta b, \qquad \ldots(4)$$

where $F=-A^{-1}\Delta A.$

Since $\| F \| = \| A^{-1}\Delta A \|$ we have $\| F \| < 1$ which implies that the matrix $(I-F)$ is invertible and

$$\|(I-F)^{-1}\| \le \frac{1}{1-\|F\|} \le \frac{1}{1-\|A^{-1}\|\,\|\Delta A\|}$$

$$= \frac{1}{1-k(A)\dfrac{\|\Delta A\|}{\|A\|}}.$$

IIT ROORKEE
NPTEL ONLINE CERTIFICATION COURSE
7

Let us see, we consider the question number 2. The question number 2; A plus delta A into x plus delta x is equal to b plus delta b. From here what we see is that, A x plus A delta x plus delta A in 2 x plus delta A into delta x is equal to b plus delta b. Now, we are given that, A x equal to b. Since A x equal to b, we have, A delta x plus delta A into x plus delta A into delta x equal to delta b. And this equation can then be written as, A times i minus F. i is the identity matrix into delta x minus equal to minus delta A into x plus delta b, which can be written as, A times i minus F A minus i A into i minus F into delta x equal to minus delta A into x of delta b, where, we have F equal to minus A inverse delta A. We can see that, you can let us put F equal to minus A inverse delta a here.

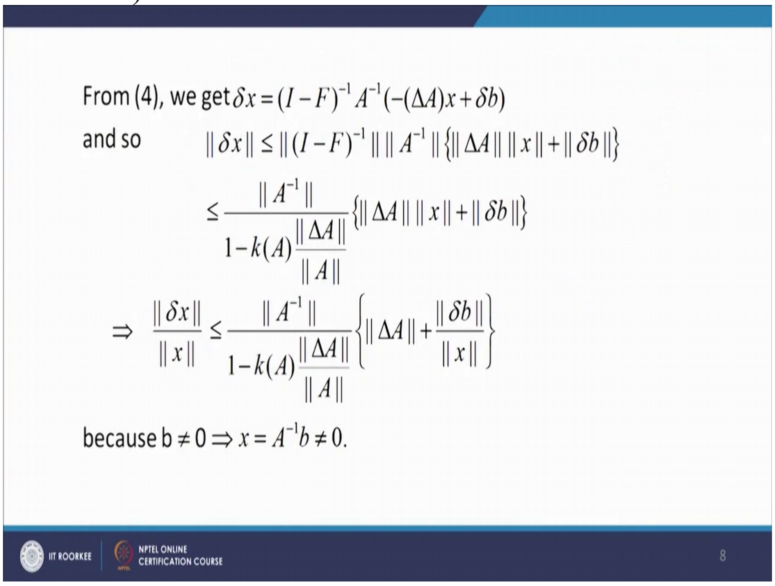
So, replacing the value of F, we get A plus A inverse delta A into delta x equal to minus delta A into x plus delta b. And you can see here, i minus A times i minus F. So, A times i plus A times i plus A inverse delta A into delta x equal to this. Now, this is identity matrix of the same order as A. So, A into i is A plus A A inverse is identity matrix. Identity matrix into delta A is delta A. So, A plus delta A into delta x equal to minus delta A into delta minus delta A into x plus delta b, which is same as this equation.

You can see we have you can bring this term to the left side, then delta A into x and then, A delta x plus delta A into delta x equal to delta b. So, this equation can be written in this form where, F is equal to minus A inverse delta A. Now, norm of F is equal to norm of A inverse delta A and we have assumed in the theorem that, norm of A inverse delta A is less than 1.

So, norm of F is less than 1 and which implies that the matrix i minus F is invertible and norm of i minus F inverse is less than or equal to 1 over 1 minus norm of F. This is a theorem which is there in the topic of matrix norms. So, it follows from there. So, norm of i minus F inverse is less than or equal to 1 over 1 minus norm of F. Now, norm of F is equal to minus A inverse delta A. So, norm of F is equal to norm of minus A inverse delta A which is equal to modulus of minus 1 into norm of A inverse delta A and this is norm of A inverse delta A, but this norm, we have assume that, the norm is the subordinate matrix norm.

So, norm of A b is less than or equal to norm of A into norm of b. So, this is less than or equal to norm of A inverse into norm of delta A. So, norm F is less than or equal to norm of A inverse into norm of delta A. So, 1 over 1 minus norm of F will be than or equal to 1 over 1 minus norm of A inverse into norm of delta A. Now, we know the condition number of A; matrix A is norm of A into norm of A inverse. So, we can write this inequality also, as norm of i minus F inverse less than or equal to 1 minus k into norm of delta A over norm of A, using where the condition number of A is equal to norm of A into norm of A inverse. From 4, we get from this equation, what do we get?

(Refer Slide Time: 12:11)



From (4), we get $\delta x = (I-F)^{-1}A^{-1}(-(\Delta A)x+\delta b)$

and so $\|\delta x\| \le \|(I-F)^{-1}\|\,\|A^{-1}\|\{\|\Delta A\|\,\|x\|+\|\delta b\|\}$

$$\le \frac{\|A^{-1}\|}{1-k(A)\frac{\|\Delta A\|}{\|A\|}}\{\|\Delta A\|\,\|x\|+\|\delta b\|\}$$

$$\Rightarrow \frac{\|\delta x\|}{\|x\|} \le \frac{\|A^{-1}\|}{1-k(A)\frac{\|\Delta A\|}{\|A\|}}\left\{\|\Delta A\|+\frac{\|\delta b\|}{\|x\|}\right\}$$

because b $\neq 0 \Rightarrow x = A^{-1}b \neq 0$.

IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE

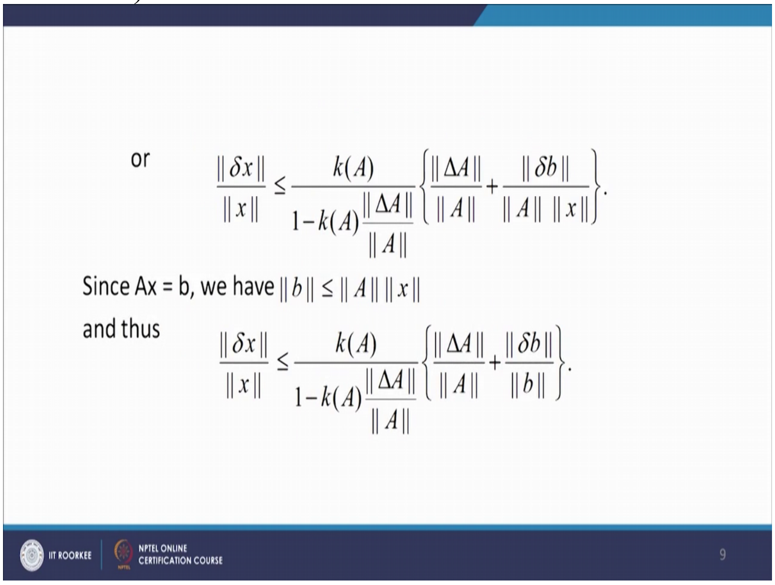
8

They we can obtain delta x as, delta x equal to i minus F inverse into A inverse into minus delta A into x plus delta b. And so, norm of delta x equal to norm of i minus F inverse into norm of A inverse into norm of minus delta A into x plus delta b, which is less than or equal to norm of delta A into norm of x plus norm of delta b (Refer Time: 12:37) which again use that this is subordinate matrix norm.

And this is less than or equal to and now, let us replace the value of i minus norm of i minus F inverse. We have seen earlier that, norm of i minus F inverse is less than or equal to 1 over 1 minus k into norm of delta A divided by norm of A. So, let us put this value here. So, this is less than or equal to norm of A inverse. So, 1 minus k into norm of delta A or norm of A multiplied by norm of delta A into norm of x plus norm of delta b. Now, we can divide this inequality y norm of x.

Because, the norm of x cannot be 0. See we know that, norm of x is equal to 0, if and only if x equal to 0. So, norm of x cannot be 0 because, x is not equal to 0. And why x is not equal to 0? Because, x is equal to A inverse b, x is equal to A inverse b and we have assumed that b is a non-0 vector. So, since b is non-0 vector, x is not equal to 0 and x is not equal to 0 implies that norm of x is strictly greater than 0. So, we can divide this inequality by norm of x and what we have norm of delta x over norm of x less than equal to norm of a inverse divided by 1 minus k A into norm of delta A over norm of A multiplied by norm of delta A plus norm of delta b over norm of x.

(Refer Slide Time: 14:31)



or
$$\frac{\|\delta x\|}{\|x\|} \le \frac{k(A)}{1-k(A)\dfrac{\|\Delta A\|}{\|A\|}}\left\{\frac{\|\Delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|A\|\,\|x\|}\right\}.$$

Since Ax = b, we have $\|b\| \le \|A\|\,\|x\|$

and thus
$$\frac{\|\delta x\|}{\|x\|} \le \frac{k(A)}{1-k(A)\dfrac{\|\Delta A\|}{\|A\|}}\left\{\frac{\|\Delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right\}.$$

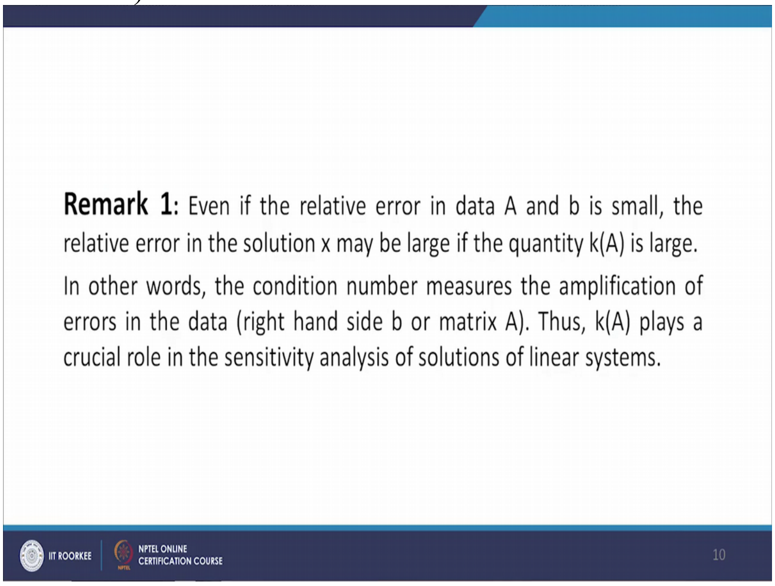
IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 9

Now, or we can write it as, norm of delta x over norm of x less than or equal to. So, again, let us replace norm of we know that, let us use this equation this equation. So, norm of A inverse is equal to k upon norm of A. When we use this, we can write it like this norm of delta x over norm of x less than or equal to k over 1 minus k into norm delta A over norm of a multiplied by norm of delta A over norm of a plus norm of delta b over norm of A into norm of x. Now, A x is equal to b A x is equal to b. So, norm of b is equal to norm of A x.

Now, this sub norm is A subordinate matrix norm. So, this is less than or equal to norm of A into norm of x. So, making use of norm of b less than or equal to norm of A into norm of x, this inequality can be further written as, norm of delta x over norm of x less than or equal to k upon 1 minus k into norm of delta A over norm of A multiplied by norm of delta A over norm of A plus norm of delta b over norm of b.

Now, thus we notice what we have proved the theorem. So, what do we notice that, the expression inside the curly bracket, norm of delta over norm of A is the relative error in A and norm of delta b over norm of delta norm of b is the relative error in b. So, even if the relative errors in A and relative error in b are small, the value of k could be large. And so, the relative error in x may also be large.

(Refer Slide Time: 16:20)

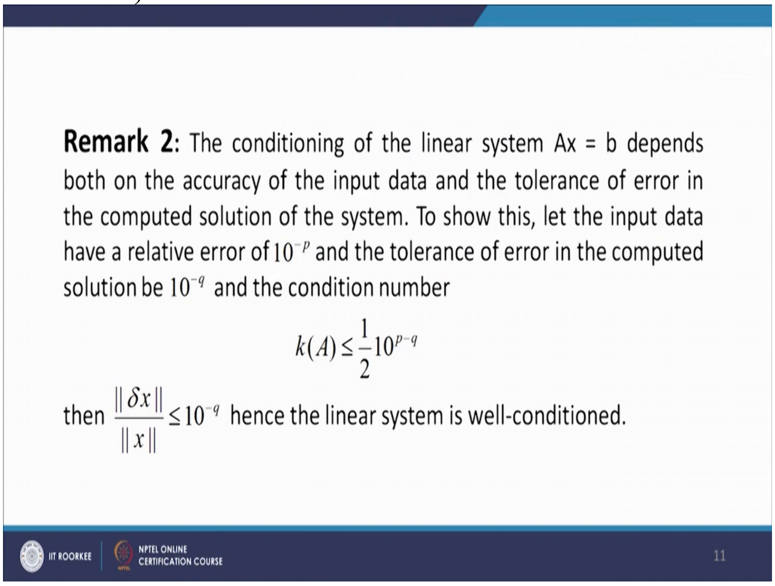


**Remark 1:** Even if the relative error in data A and b is small, the relative error in the solution x may be large if the quantity k(A) is large. In other words, the condition number measures the amplification of errors in the data (right hand side b or matrix A). Thus, k(A) plays a crucial role in the sensitivity analysis of solutions of linear systems.

IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 10

So, what we see that, even if the relative error in delta A and b is small, the relative error in the solution, x maybe large if the quantity k is large. In other words, the condition number measures the amplification of errors in the data right hand side b or matrix A. Thus, k A plays a crucial role in the sensitivity analysis of solutions of linear systems.

(Refer Slide Time: 16:51)



**Remark 2:** The conditioning of the linear system Ax = b depends both on the accuracy of the input data and the tolerance of error in the computed solution of the system. To show this, let the input data have a relative error of $10^{-p}$ and the tolerance of error in the computed solution be $10^{-q}$ and the condition number

$$k(A) \le \frac{1}{2} 10^{p-q}$$

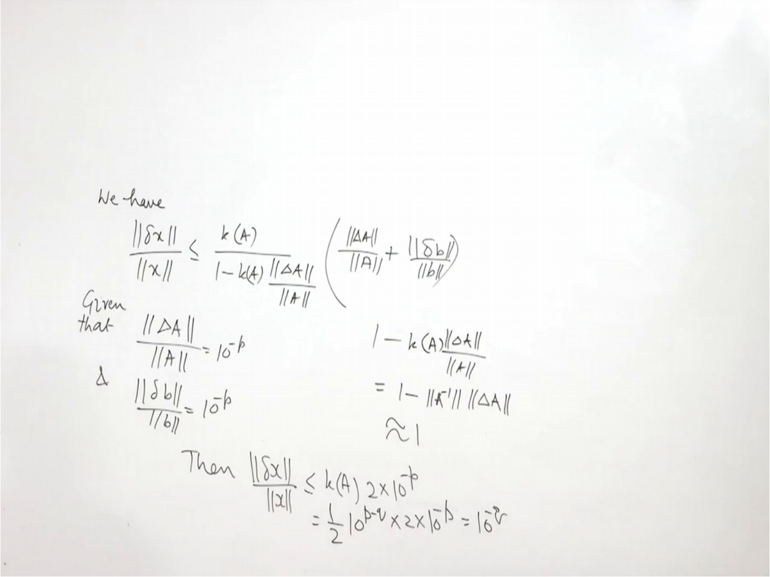
then $\frac{\|\delta x\|}{\|x\|} \le 10^{-q}$ hence the linear system is well-conditioned.

IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 11

Now, the conditioning of the linear system A x equal to b depends both on the accuracy of the input data and tolerance of error in the computed solution of the system. We can this let us say, we have the input data, which is the relative error of 10 to the power minus p and the tolerance of error in the computed solution b 10 to the power minus q and let us suppose that, the condition number is such that, k is less than or equal to 1 by 2 times 10 to the power minus p minus q.

Then, norm of delta x over norm of x is less than or equal to 10 to the power minus 2 q. And so, we can say that, the linear system is well conditioned. Now, let us see how we get this inequality.

(Refer Slide Time: 18:00)



We have

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{k(A)}{1 - k(A)\frac{\|\Delta A\|}{\|A\|}}\left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|}\right)$$

Given that $\frac{\|\Delta A\|}{\|A\|} = 10^{-p}$

& $\frac{\|\delta b\|}{\|b\|} = 10^{-p}$

$$1 - k(A)\frac{\|\Delta A\|}{\|A\|} = 1 - \|A^{-1}\|\,\|\Delta A\| \approx 1$$

Then $\frac{\|\delta x\|}{\|x\|} \leq k(A)\, 2 \times 10^{-p}$

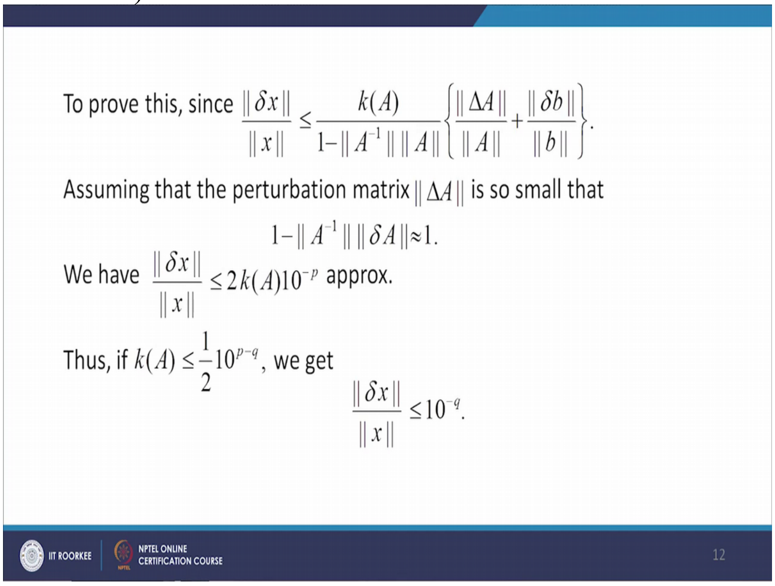
$$= \frac{1}{2} 10^{p-q} \times 2 \times 10^{-p} = 10^{-q}$$

So, we have norm of delta x over norm of x less than or equal to k upon 1 minus k A times norm of delta A over norm of A into norm of delta A over norm of A plus norm of norm of delta b over norm of b. This is what we have with us. Now, let us see, we have assume that the input data has a relative error of 10 to the power minus p. So, we are given that norm of delta A over norm of A. This is equal to 10 to the power minus p and norm of delta b over norm of b.

This is i also equal to 10 to the power minus p. Both of them have a relative error of 10 to the power minus p. And then, let us assume that here, k A is equal to norm of A into norm of A inverse. So, this is 1 minus norm of A inverse into norm of delta A. So, let us assume that norm of delta A; that is, the perturbation matrix in the perturbation matrix is so small that, we can assume that this 1 minus k A times norm of delta A divided by norm of A. This is equal to 1 minus norm of A inverse into norm of delta A. Let us assume that it is approximately equal to 1.

So, we assume that, the perturbation matrix norm of delta A is so small that, 1 minus norm of A inverse into norm of delta A is approximately equal to 1. Then, norm of delta x over norm of x will be less than or equal to this is 10 to the power minus p. So, 2 times this is to this is a proximity to tie less than or equal to k A times 2 into 10 to the power minus p and k we are assuming to be 1 by 2 10 to the power p minus q. So, this into 2 into 10 to the power minus p and this is equal to 10 to the power minus q.

(Refer Slide Time: 21:12)



To prove this, since $\frac{\|\delta x\|}{\|x\|} \le \frac{k(A)}{1-\|A^{-1}\|\,\|A\|}\left\{\frac{\|\Delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right\}.$

Assuming that the perturbation matrix $\|\Delta A\|$ is so small that

$$1-\|A^{-1}\|\,\|\delta A\|\approx 1.$$

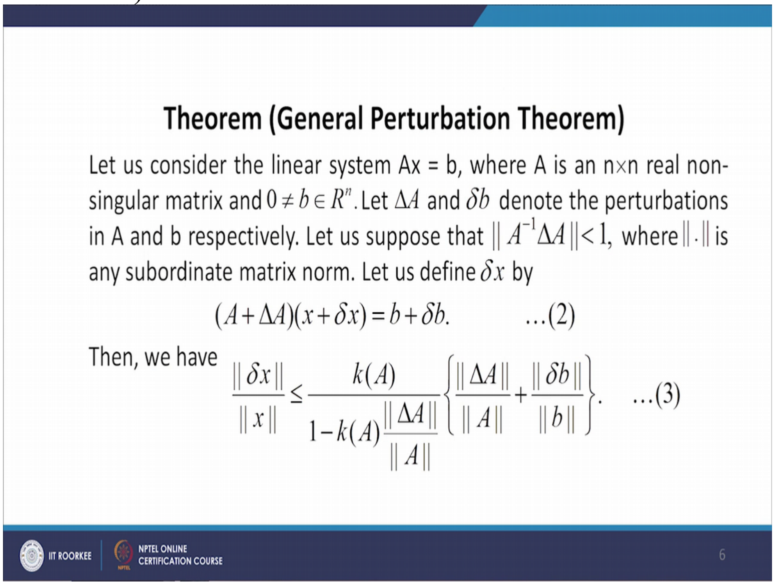
We have $\frac{\|\delta x\|}{\|x\|} \le 2k(A)10^{-p}$ approx.

Thus, if $k(A) \le \frac{1}{2}10^{p-q}$, we get

$$\frac{\|\delta x\|}{\|x\|} \le 10^{-q}.$$

So, we have the relative error in x is less than or equal to 10 to the power minus q and which is a small. So, we can say that, the system is in good condition; I mean well-conditioned.

(Refer Slide Time: 21:25)



### Theorem (General Perturbation Theorem)

Let us consider the linear system Ax = b, where A is an n×n real non-singular matrix and $0 \ne b \in R^n$. Let $\Delta A$ and $\delta b$ denote the perturbations in A and b respectively. Let us suppose that $\|A^{-1}\Delta A\| < 1$, where $\|\cdot\|$ is any subordinate matrix norm. Let us define $\delta x$ by

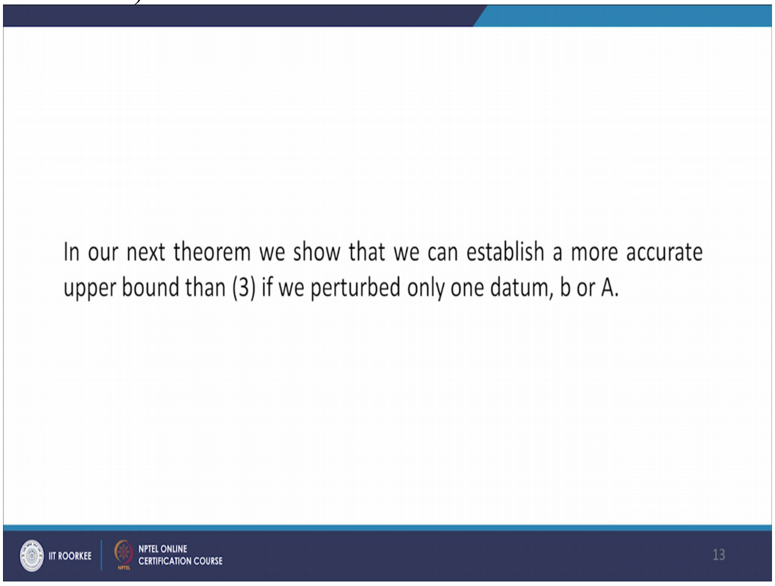
$$(A+\Delta A)(x+\delta x) = b+\delta b. \qquad \ldots(2)$$

Then, we have $$\frac{\|\delta x\|}{\|x\|} \le \frac{k(A)}{1-k(A)\frac{\|\Delta A\|}{\|A\|}}\left\{\frac{\|\Delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right\}. \qquad \ldots(3)$$

Now, in our next theorem, we shall show that, we can establish a more accurate upper bound then the bound 3 we want 3 is this one norm of delta x over norm of x less than or equal to norm k A time upon 1 minus k A into norm of delta A norm of A into norm of delta A over norm of b plus norm of delta b over norm of b. And we shall see that we can get a better upper bound and this, there we will assume that, either there is error in b or there is error in A.

So, then we will see that, if there is perturbation in A, only in A, then we get a better, we get a result which is optimal. Also, when there is only perturbation in b, we get the result which is again optimal. So, we can say that we get a better result than this.

(Refer Slide Time: 22:14)



In our next theorem we show that we can establish a more accurate upper bound than (3) if we perturbed only one datum, b or A.

IIT ROORKEE
NPTEL ONLINE CERTIFICATION COURSE
13

So, that is our aim to which we will show in the theorem 1. In our next theorem, we show that, we can establish a more accurate upper bound than 3, if we perturbed only one datum that is b or A. With this, I would conclude my lecture.

Thank you very much for attention.