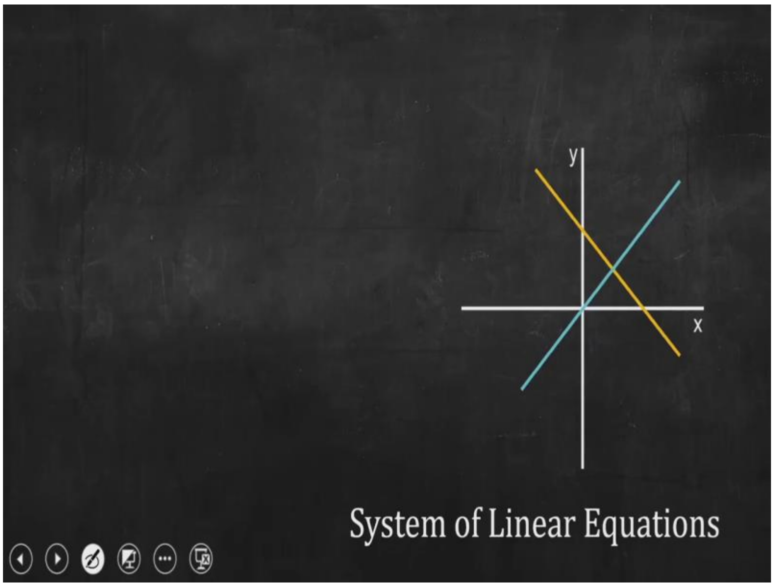
Data Analysis for Biologists

Professor Biplab Bose
Department of Biosciences & Bioengineering
Mehta Family School of Data Science & Artificial Intelligence
Indian Institute of Technology, Guwahati

Lecture 12

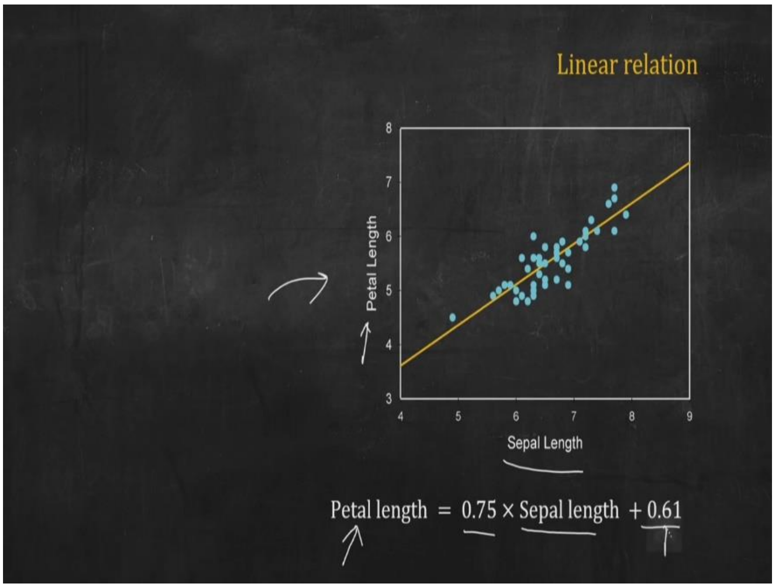
System of Linear Equations

(Refer Slide Time: 0:30)



Hello everyone, we are discussing linear algebra for our data analysis course. And in this lecture, I will discuss about system of linear equations and how to solve them.

(Refer Slide Time: 0:52)

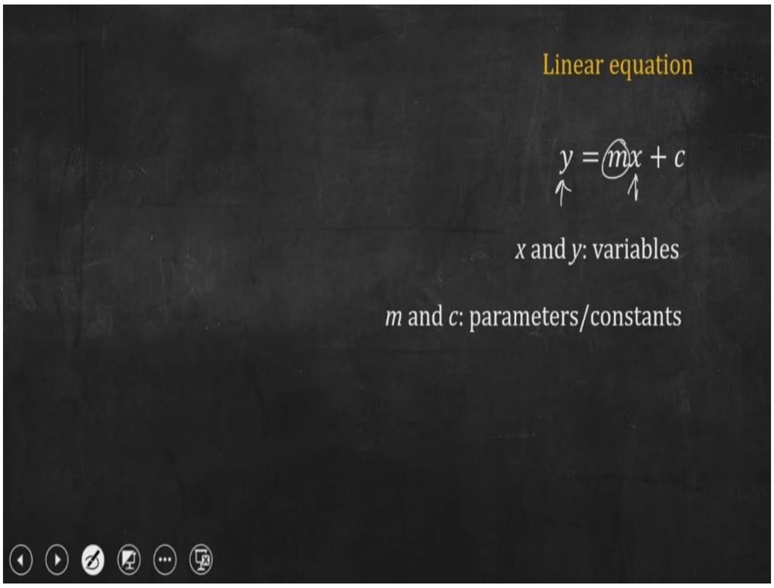


So, let me start with a data set which is very popular among the people who learn machine learning and data analysis this is called Iris data set, it has been collected by one botanist and it has been used by a famous statistician Fisher in his old paper, I think, in 1936 or something for some statistical work and it has the data for the size of the morphological properties of certain Iris flowers.

And what I have done is a large data set from that I have picked up a particular part and I have plotted the data here, you can see the diagram. So, I have sepal length in the horizontal axis and petal length in the vertical axis, just look at the data. Looking at the data, you can easily see there is a linear trend, is not it? There is a linear trend between the sepal length and the petal length.

So, your obvious tendency would be to why do not I fit a straight line to that, and you must have done that this type of exercise in many times using an excel sheet, you may have done fitting a trend line which is a linear one. And I have just done that and in our technical jargon we call it I have done a linear regression, we will learn linear regression in it later on. So, what I have got I have got a real relation between sepal length and petal length and that is written here, petal length is equal to 0.75 into sepal length plus a constant term. So, this is typical of a linear equation, is not it?

(Refer Slide Time: 2:20)

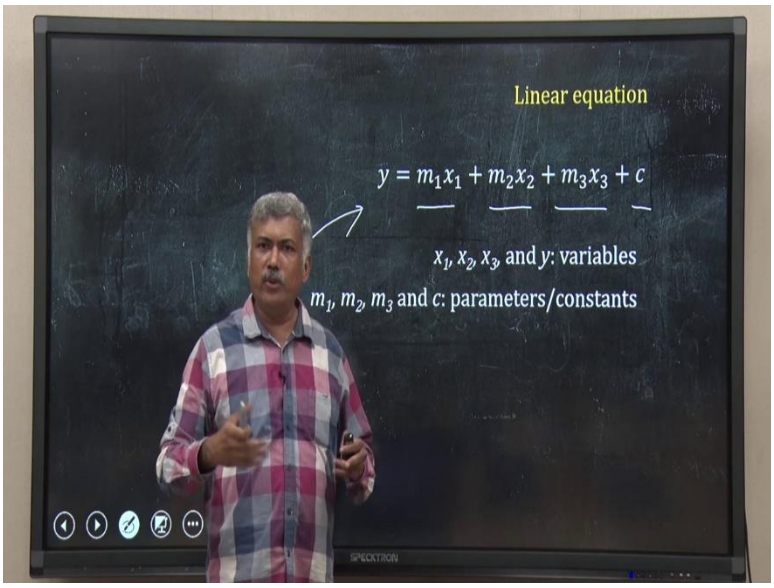


In our school, we have learned about equation of a straight line, which is a linear equation

$$y = mx + c,$$

I have two variable y and x and m and c are constants. Similarly, there is nothing specific that I have to keep only two variable y and x, I can have multiple variable, but I can still have a linear relation.

(Refer Slide Time: 2:44)

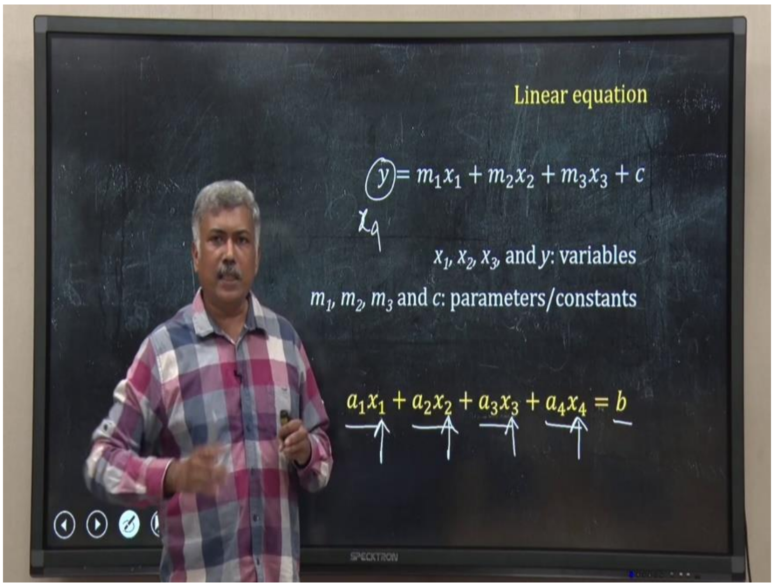


For example, I can have something like this,

$$y = m_1x_1 + m_2x_2 + m_3x_3 + c$$

This one is also linear, because you look at it, the power of these x and y's is 1 and there is no multiplication between these terms this variable. So, this is a linear equation representing linear relation between these variable three x's and one y, whereas m1, m2, m3, and c are constants. Usually in our data analysis and linear algebra, we do not write this equation in this way, this is the way you usually write when you do coordinate geometry in school.

(Refer Slide Time: 3:23)

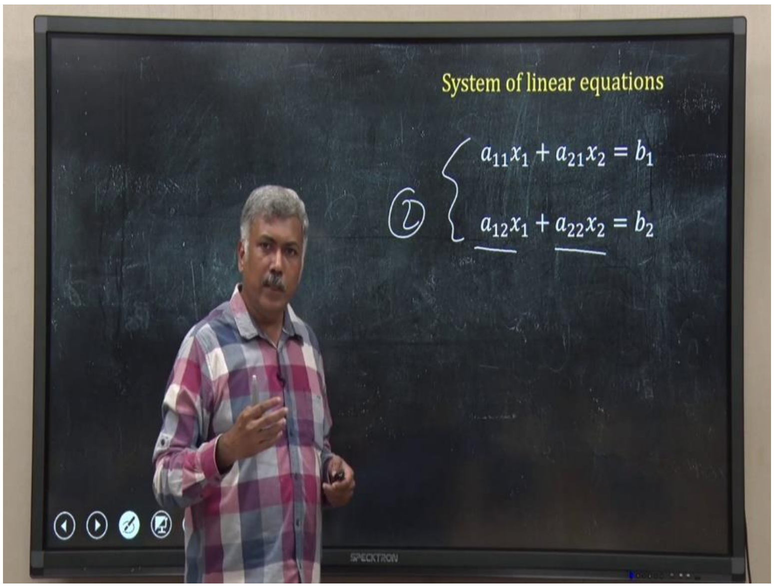


What we do, we write the same thing in this fashion,

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = b$$

, where x1 x2 x3 and x4 are my variables, you can simply replace this y by x4 and then rearrange the m and c term you will get back this one a1 x1 plus a2 x2 plus a3 x3 plus a4 x4 equal to b and this is a linear equation because all these variable has power 1 and there is no fancy function like sine exponential or something like that, or there is no product between these variables. So, this is a linear equation. Now, this is only one linear equation, but in our data analysis, usually we will have multiple linear equation.

(Refer Slide Time: 4:11)



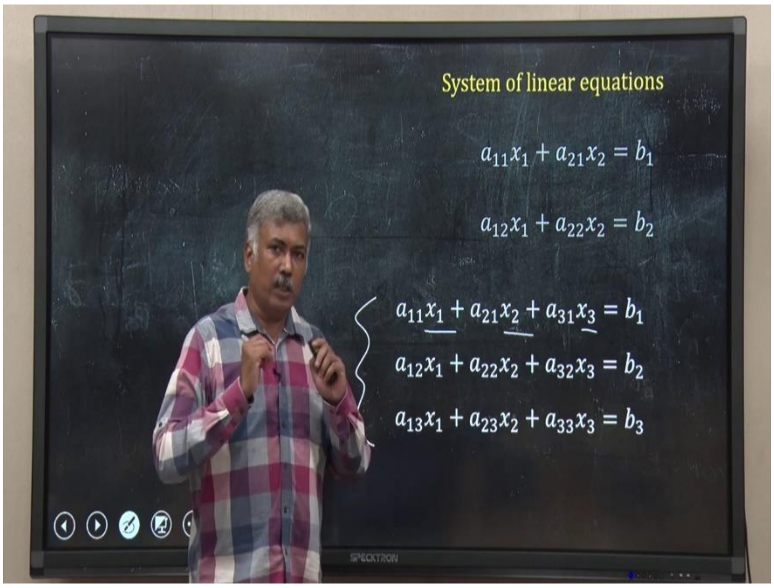
So, what do I get? I get a system of linear equations and I have shown an example here. So, I have two equation here, so, this is a system of two linear equations,

$$a_{11}x_1 + a_{21}x_2 = b_1$$

$$a_{12}x_1 + a_{22}x_2 = b_2$$

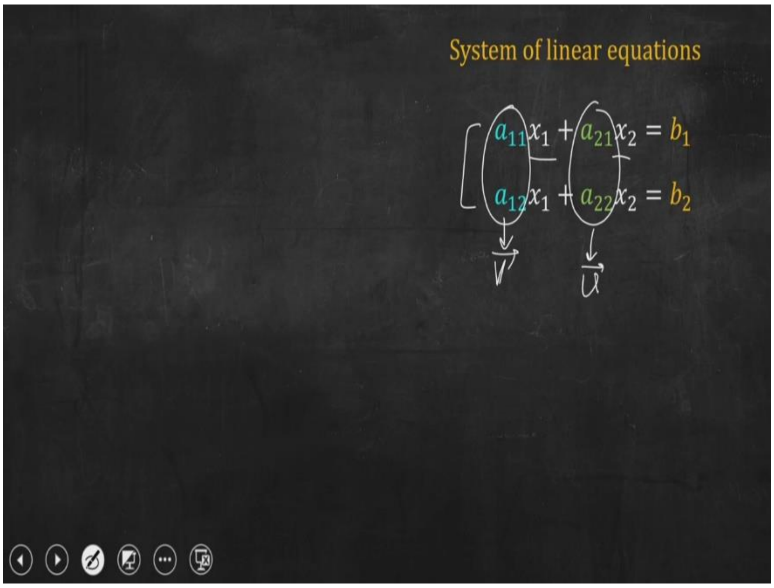
Notice that in both the equation I have x1 and x2 these a11 a22 these are the constant and they can be any value, negative, positive, zero whatever it is. Similarly, b1 and b2 also, I can have a system of equations involving three variable.

(Refer Slide Time: 4:51)



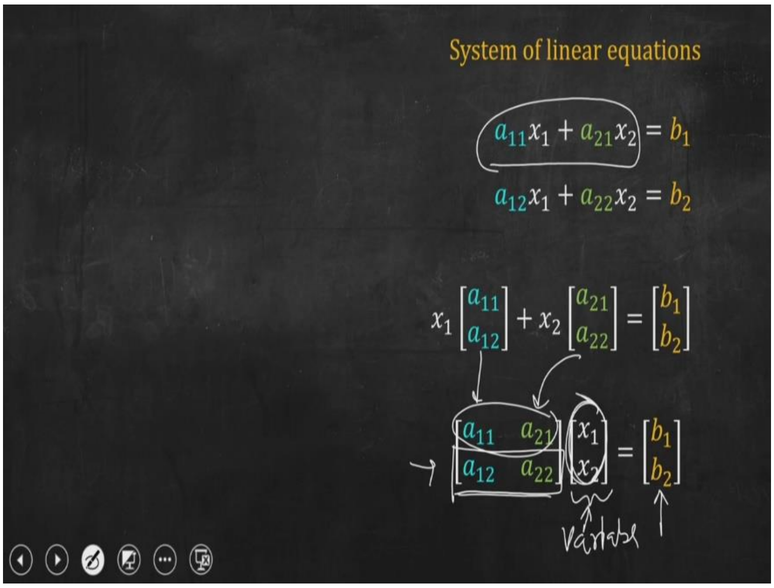
For example, I have shown here, three variable based system of equation I have three variables, x 1 x2 x3 and I have written three equation, I have a purpose for writing three equation, I could have written five equation also, it does not matter, but as a whole, I have multiple equations involving these variables that is why it is called system of linear equations, all of them are linear, so it is a system of linear equations.

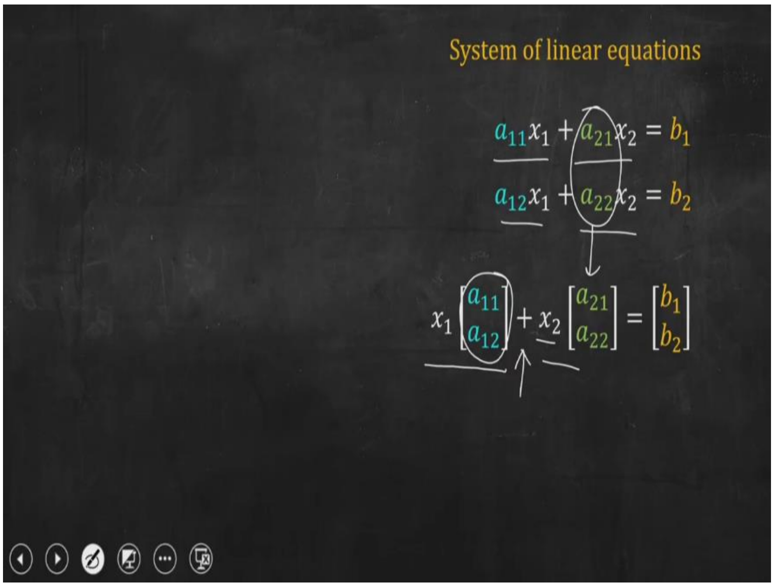
(Refer Slide Time: 5:27)



Now, let me move from this system of equation to a matrix based on, matrix or vector-based representation of a system of linear equations. How do I do that? Let me take a two equation, two variable system. I have two variable x1 and x2 and I have two equations. Take the first constant, which are multiplied with x1 and take the second one, convert this one in one vector, suppose V and make this one another vector. So, I take this constants, or parameters associated with each of these variable write together and I create a vector.

(Refer Slide Time: 6:03)





$$x_1\,[a_{11}\ a_{12}] + x_2\,[a_{21}\ a_{22}] = [b_1\ b_2]$$

$$\begin{vmatrix} a11 & a21 \\ a12 & a22 \end{vmatrix} \quad \begin{vmatrix} x1 \\ x2 \end{vmatrix}$$

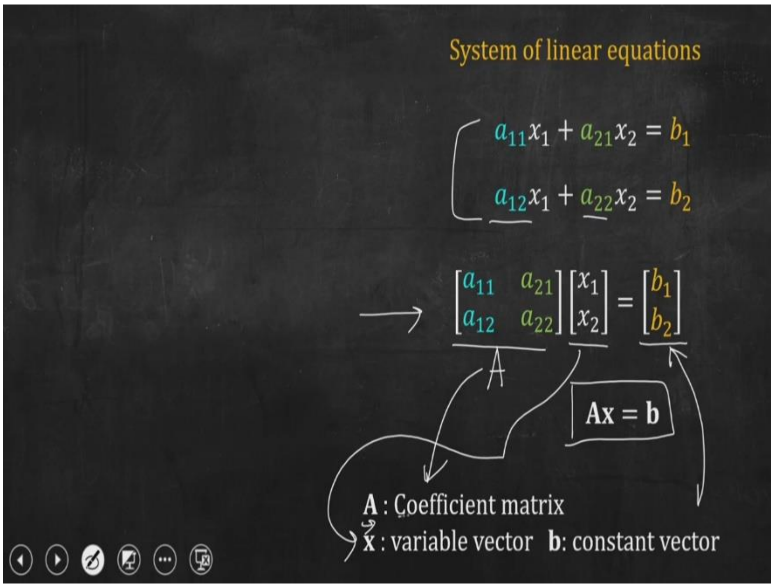
$$= [b_1\ b_2]$$

So, how I have done I can, how do I go, what do I get out of it, I get something like that. So now I can multiply this first vector, a11 a12 with x1, then I will get back the same thing plus I multiply x2 with the second vector that I have created. So, I will get back these terms and the constant on the right-hand side, I represent also as a column vector, b1, b2. Now, using the rules of matrix multiplication, you can easily represent this relationship as this format.

So, what I have? I have a matrix here, which has these vectors, arranged side by side as column vectors of that matrix, and then I multiply that with a vector made of my variables. So, this is a variable vector. And obviously, I have the right-hand side constant as a separate column vector. Just try to remember the rules of multiplication, what I have to do? I have to take this row and multiply with this one.

So, then I will get back this part of the first equation a11 into x1 plus a21 into x2. Now, take the second row of this matrix and multiply with the same vector x1 x2 what will happen, you will get the second equations left hand side a12 x1 plus a22 x2. So, that means a system of linear equation is given to me, I can represent it as a compact form using a matrix and few vectors.

(Refer Slide Time: 7:47)

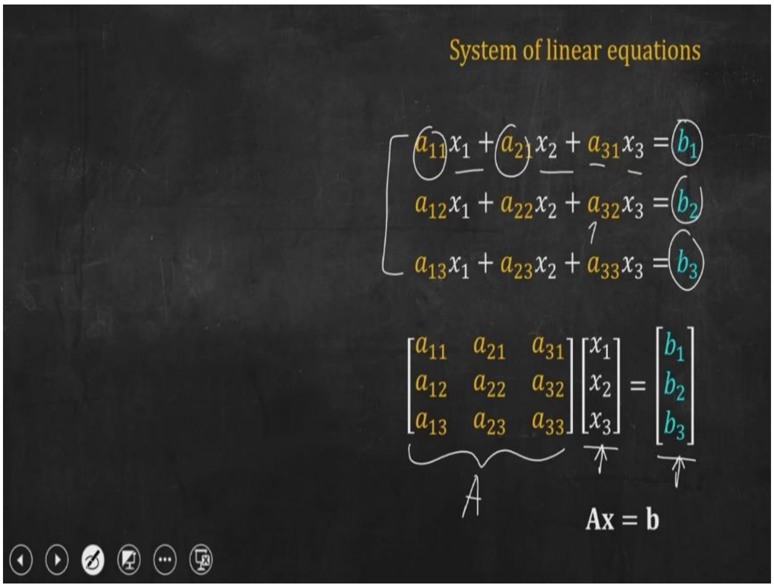


So, as a whole, I have written it down in a shortened form. So, this is the given equation for me, and I can represent that as a matrix into a vector is equal to another vector. So, this first matrix is called A and it is called the coefficient matrix, whereas this one is called the X vector, which is the variable vector and this one is my constant vector. So, what I have got? I have got,

$$Ax = b$$

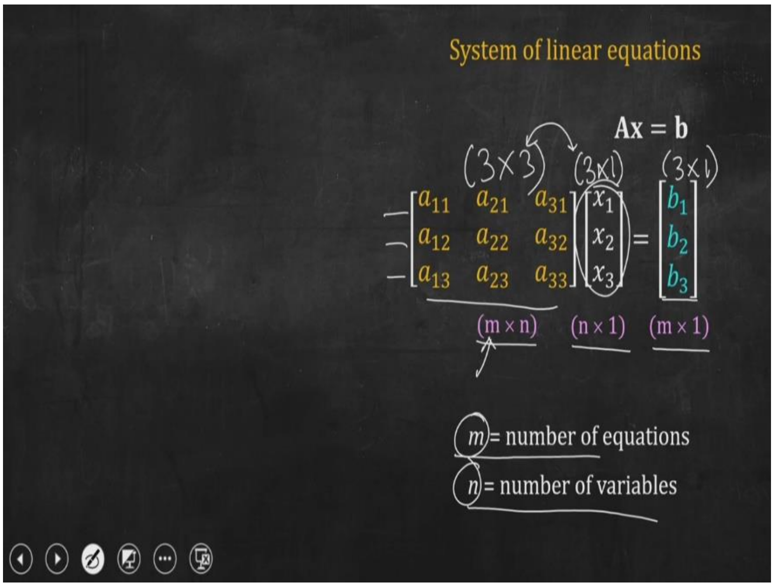
This is the shorthand form of writing a system of linear equations. So, A has all the constants, so it is called coefficient matrix. So, I have a coefficient matrix into the vector for variable, here in case of variable x_1 and x_2, this multiplication product is equal to the vector on the right hand side which is b. So, Ax equal to b.

(Refer Slide Time: 8:55)



Now I can do the same thing if I have three equations. So, I have three equations here for three variables. So, I will take the constant term which are marked in yellow, you can see easily, and I will create the coefficient matrix A. Whereas the variable vector, the column vector will be x1 x2 x3 stacked one over each and on the right-hand side of the equal sign I will have the vector representing the right hand side term of my original linear equations. So again, I will get back Ax = b. So, this is a shorthand form of writing a system of linear equations.

(Refer Slide Time: 9:40)



Now, when we are writing this one, you must have noticed one thing, the rule of multiplication must be satisfied. So, what is the dimension of this one? This is 3 into 3, 3 rows and 3 columns.

What is the next vector? Vector is three into one, because I have three rows, so obviously, this is equal to this one so, I can do the multiplication and the resultant will be 3 into 1 vector and I have 3 into 1 vector here, so the rule of multiplication is satisfied.

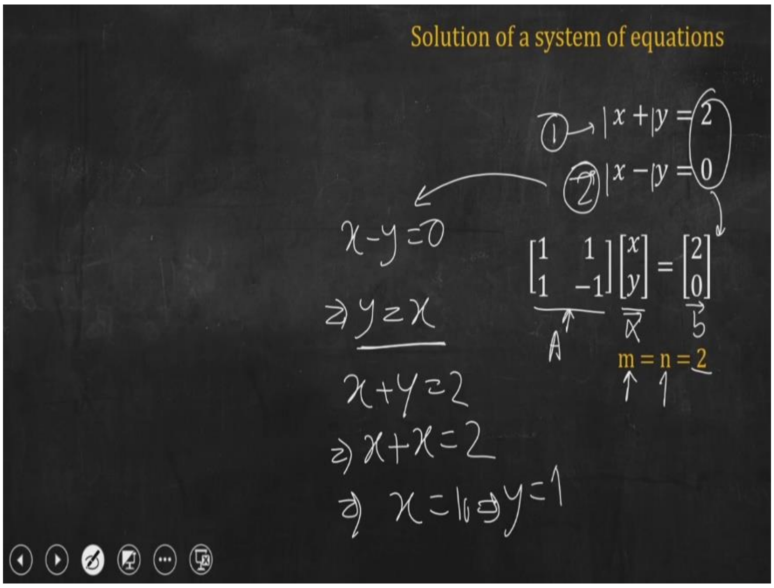
So, usually, this m here I have three equations. So, number of equation is represented by m, it is not mandatory that you have to call it m for our discussion we will call now number of equation equals m and the number of variable in this case is 3 that is n. So, the coefficient matrix is m into n matrix, whereas, the variable vector is n into 1 vector and the constant vector on the right hand side is m into 1 vector. This way of numbering in the, naming the number of equation and number of variables is important because, I will enter into discussion where the, if we will see the effect of m and n on our calculations.

So, we have got a shorthand representation or rather a linear algebra-based representation of a system of linear equations. But why I am taking so much of trouble of that? Because many times we will use linear algebra to solve a system of linear equations. What do I mean by solve a system of linear equations? We will have some data and that data may be represented by this matrix and this vector and I want to know the value of x1 x2 and x3.

So, in my problem in my data analysts the coefficient matrix and that constant vector these may be coming from some equation or from data experimental data, whereas, the variable x1 x2 x3 will be unknown to me, so, I have to calculate these unknown variables. So, that is why when it is solving linear equation, we call these variables unknown.

So, we have to solve this system of equation and linear algebra helps us immensely to solve large system of linear equations, and most of the solver in your computer will use linear algebra to solve it. So, let us take a small system and try to solve that one using first, general method and eventually using linear algebra base method.

(Refer Slide Time: 12:20)



So, I will take an example, very simple,

$$x + y = 2$$

$$x - y = 0$$

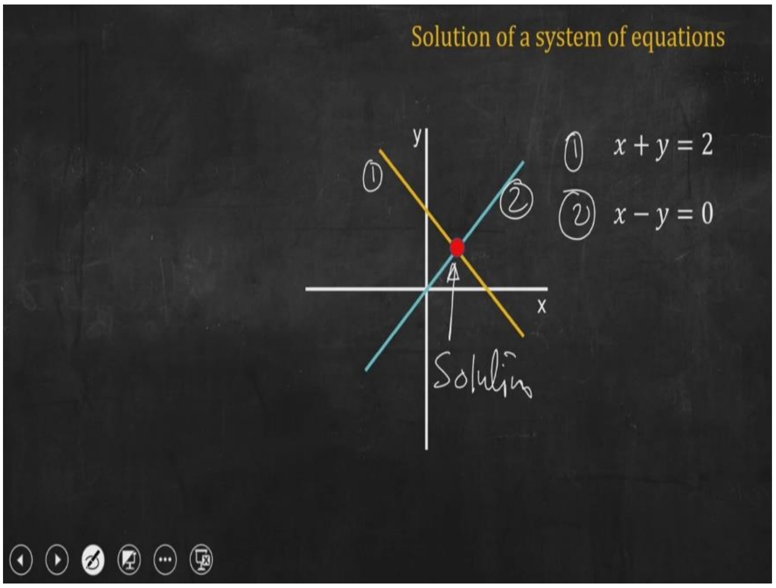
I can represent it as a coefficient matrix in linear algebra, I have,

1	1		x	=	2
1	-1		y		0

1 here 1 here 1 here 1 here. So, my A, this is my A, A will be 1 1 1 minus 1 and this is my X vector and this is my b vector, 0 and 2 this is. The number of equation is m, number of variable is 2, how will you solve it? Very easy to solve every one of us has learned that in school, so, what I will do, I will take the second equation and from that, I will get what? I will get x minus y equal to 0 that means y equal to x.

Now, take the first equation and put that x plus y equal to 2 that means, I will replace x, y. So, this will be x, y is replaced by x as I have done got from the second equation that is 2. So, that means x equal to 1, so, that will give me y equal to 1 also. So, 1 1 is the solution, as simple as that you do not need linear algebra or computer to do that.

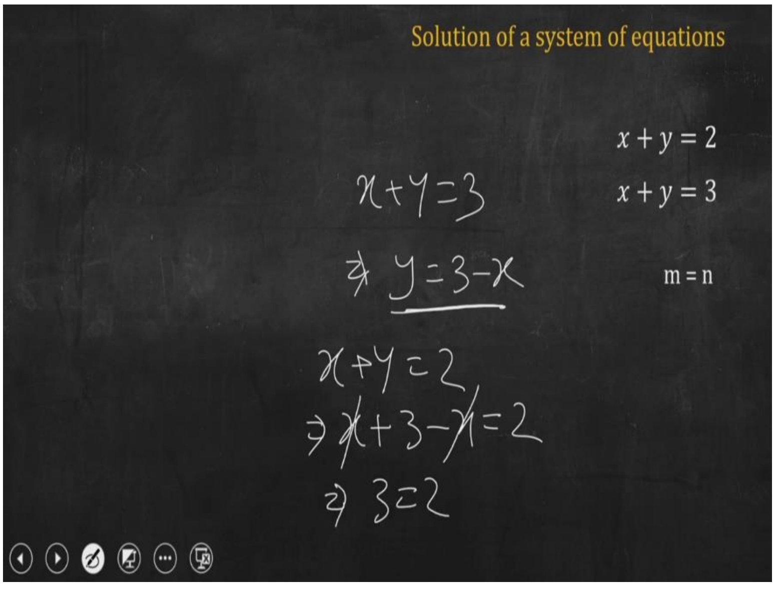
(Refer Slide Time: 13:53)



But let us look into the geometric meaning of this solution. So, what I have done, this is the first equation, this is the second equation. In this 2D space, I have drawn those two, these are straight line, these are linear equation that means they are straight line. And they are intersecting at this red dot that is the solution, because at that red dot, the value of x and y satisfies both the line straight line. So, that is my solution.

So, for any linear system of equations, the intersection of these lines in my space will be the solution. So, we will try to use this concept to explore some system of linear equations, all of them are simple, but they will have different behavior.

(Refer Slide Time: 14:40)



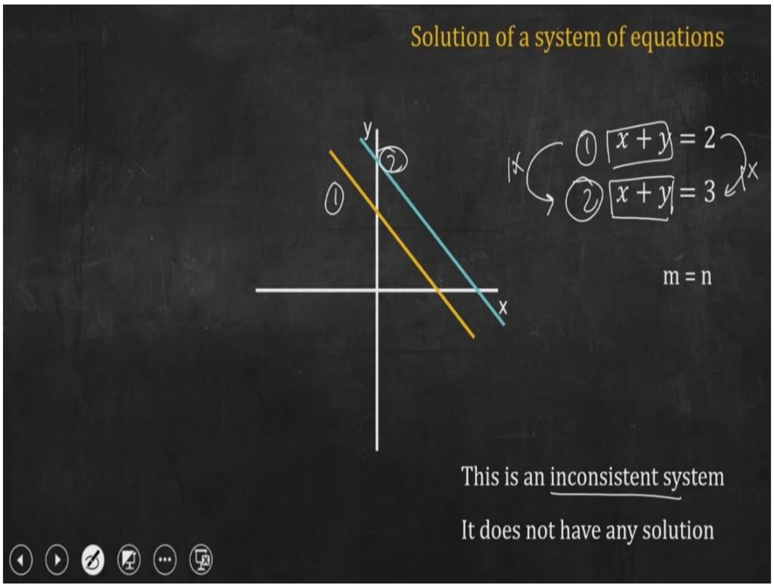
Let us take the first 1,

$$x + y = 2$$

$$x + y = 3$$

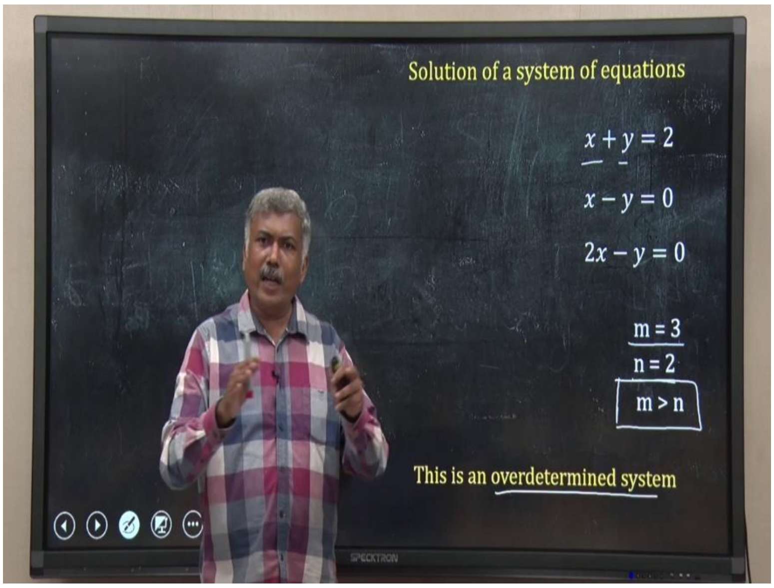
just try to solve it, the second equation will give me x plus y equal to 3 that means y equal to 3 minus x. When I take that and put in the first equation, first equation x plus y equal to 2, I will replace y, x plus 3 minus x from here from this second solution, second equation. So, that is equal to, x and x get cancelled, I get 3 = 2 absurd this can be possible, you are getting an absurd answer. Why is it so? Let us plot the graph and it will be clear why it is so.

(Refer Slide Time: 15:26)



This is these are the two equations and I have the two plots, this is possibly one, and this is the second one, you can see these two lines are parallel, these two straight lines are parallel to each other, so they do not intersect, so there is actually no solution to this. In fact, this system of equation is called the inconsistent system, because if you notice the equation, this one the first one can be multiplied with one to get the same on the left-hand side. But on the right-hand side, if I multiply one, I do not get the same thing that means, there is an inconsistency and the left hand side and the right hand side of these two equations. So, that is why it is called a inconsistent system and it does not have any solution, that means all linear system of equations will not have solution, we have just started to explore those.

(Refer Slide Time: 16:33)



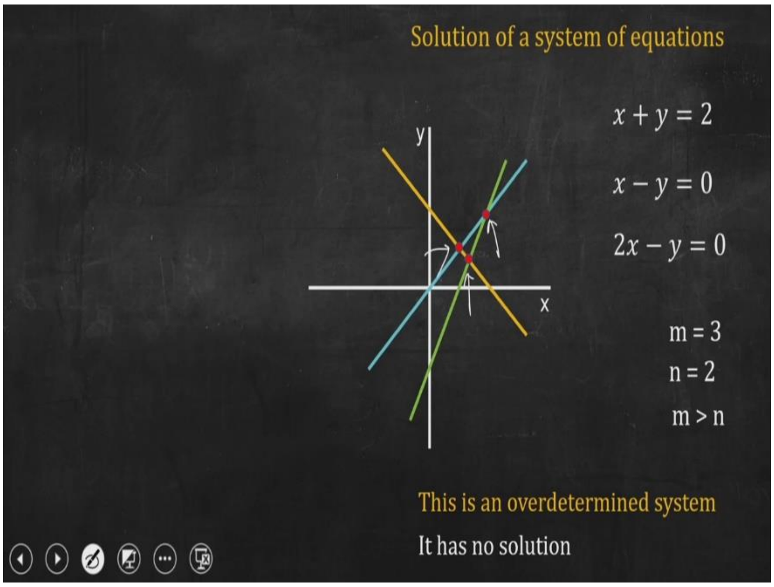
$$x + y = 2$$

$$x - y = 0$$

$$2x - y = 0$$

Now, let me take another example. Now, I have increased the number of equations. So, my m is 3, number of equation is three, but how many variables are there two, x and y. So, n equal to 2, that means my m is greater than n, number of equation is more than number of variable or number of unknowns. So, I have more equations than unknown. So, this system is called over determined system. I am not going to trying to solve this equation manually, what I will do? I will graphically represent this system three linear, three lines.

(Refer Slide Time: 17:09)

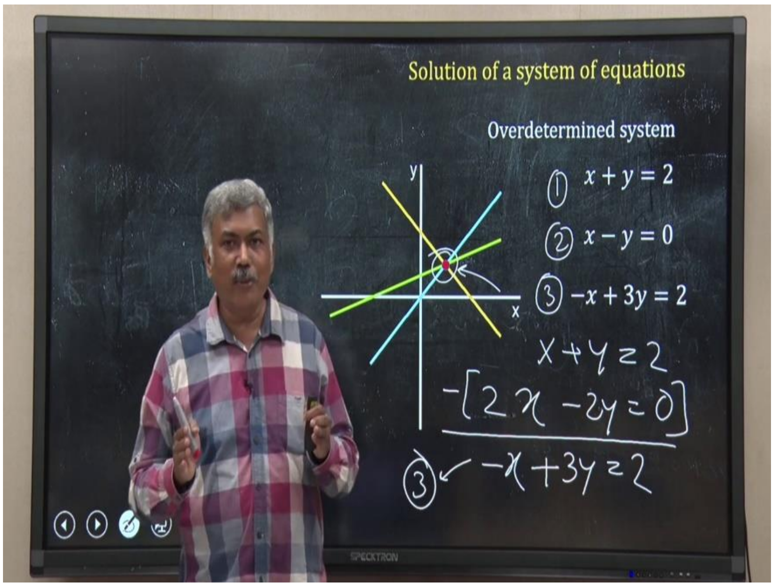


So, this is what we get, we have three equations and they does not, they all these lines do not intersect at one single point what they have? They have three intersections and none of these three is common to all of these three, that means this system has no unique solution. So usually, usually overdetermined system in general will not have any solution. As I said, in over determined system my number of equations will be more than number of variables or unknown, and they will not have any solution as I am showing in this example.

So, you will be pretending to avoid them in your data analysis. But unfortunately, for most of the cases, in data analysis we will very frequently encountered this type of overdetermined system, where the size of your data will be much bigger than the number of variables. So, your A matrix will be a rectangular one, a tall and skinny one.

So, is there a way to handle this? Yes, we'll not be able to find a unique solution, as I explained using this diagram, but we can use some optimization technique to find the optimum solution. And that is all done in case of linear regression, that we will learn later in this course, we try to find out it optimum solution for this overdetermined system as it does not have a unique solution.

(Refer Slide Time: 18:47)

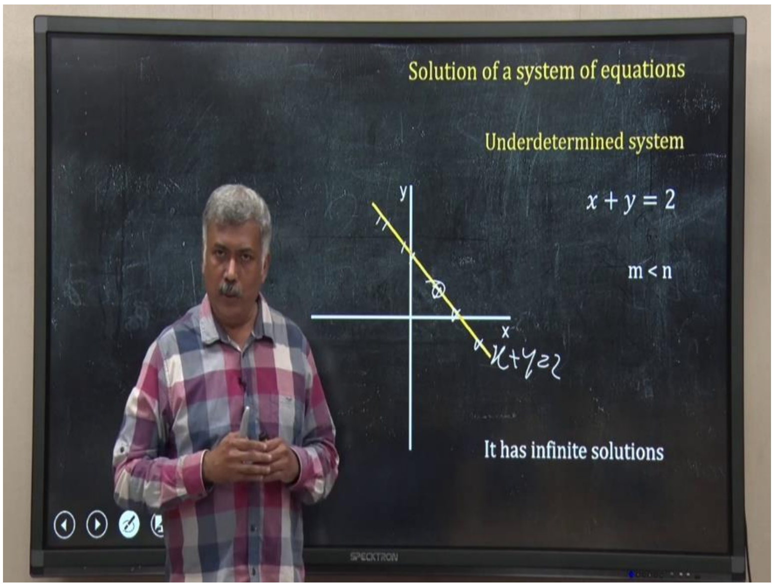


Now, let me give an example of another over determined system. Here also three equation two unknown, x and y. Let us plot the line. Now, in this case three equations are intersecting at one unique point that means I have the unique solution. How can that be possible? Just now few seconds back I said if I have an overdetermined system, when number of equation is more than number of unknown variable will not have solution and I have shown the example, then how come this is happening? Where we are going wrong? Just pay attention to the equation and you realize where is the problem?

If I take this first equation, let me mark them 1, 2 and 3. So, I take the first equation x + y = 2. And then I multiply the second equation with 2, I get 2x - 2y = 0. And I subtract this whole thing from the first one. So, what do I get? x - 2x which is -x, y + 2y it will be +3y and 2 - 0 is 2. I have got the third equation. That means the third equation is not independent, I can represent it by combining linearly, the equation 1 and 2. That means these three equations are not linearly independent.

So, I can actually simply replace, remove this third equation. And I will get back to a situation where I have two equation and two variable and I have a unique solution. So, in this case is overdetermined, but as the equation are not linearly independent, we can reduce the system, and we have the solution. Now, let me move to the opposite side, we have a overdetermined, now I will move to underdetermined system, what is that?

(Refer Slide Time: 20:49)



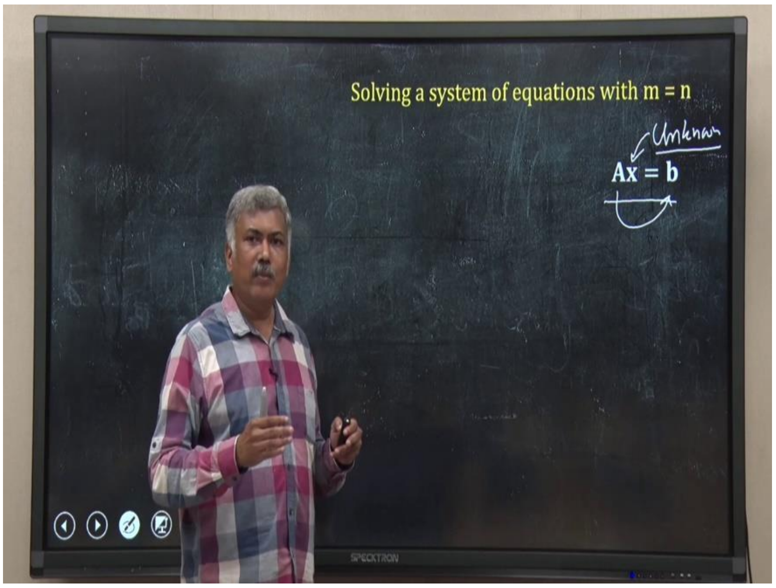
I have given you a simple one equation x + y = 2. That means number of equations m is less than number of variable n = 2. What will be the solution of that? You can easily guess that any value of x and y possibly will solve this and in fact that is true any value, I have in finite solution, I have drawn the diagram, this is my equation, this is x + y = 2, any point, any point on this line will satisfy my equation.

That means for underdetermined system, where the number of equation is less than unknown, or the number of variable the system is underdetermined, and have infinite solution, just like overdetermined system, I do not have any unique solution here at all, rather I have infinite solution, all of them are solution and that is not useful for us.

So, but unfortunately, in many data analysis cases, you will encounter this, although you may not discuss this one in our course, what do people do in that case? They again use some sort of optimization where they put some extra constraint on what is the possible value of x and what is the possible value of y, so that they can actually zoom into a particular region of this and find the optimal solution.

So, what I have discussed till now is that, if I have a linear system of equation, if it is overdetermined, I do not have a solution, rather we usually use optimization-based method to find the optimum solution. Whereas if underdetermined, I have infinite number of solution, again that is not very useful. So, the useful thing is when m = n. And let us see how linear algebra can be used when m equal to n, number of equation is equal to a number of variables, how we solve it?

(Refer Slide Time: 22:37)

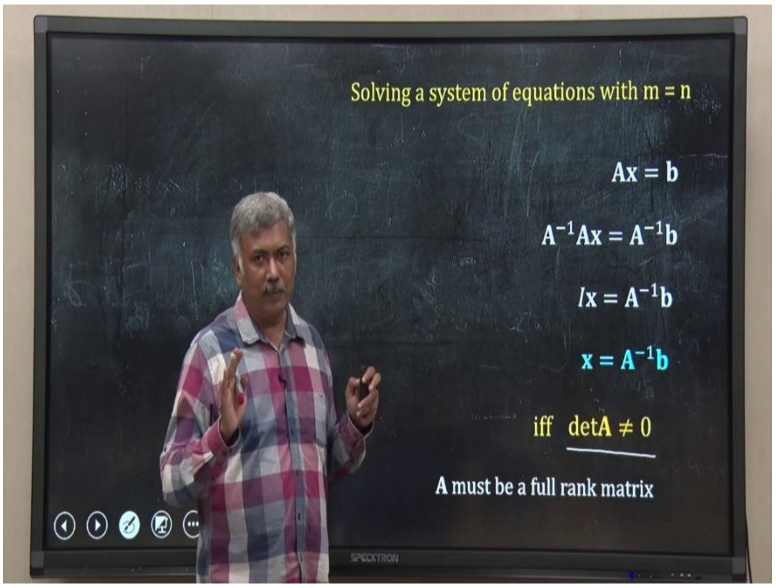


So again, I will represent the whole system of equation as in a matrix notation,

$$Ax = b$$

A is the coefficient matrix, X is the variable vector, b is the constant vector. Now, I want to know X, this is my unknown, this is my unknown. So, I want to know X. Will it not be nice if I can simply take A on the side and divide b by A I will be left with X on the left hand side, I am separating, if I can do that. But unfortunately, I cannot do division by a matrix. I cannot divide a vector by a matrix.

(Refer Slide Time: 23:25)



$$Ax = b$$

$$A^{-1}Ax = A^{-1}b$$

$$Ix = A^{-1}b$$

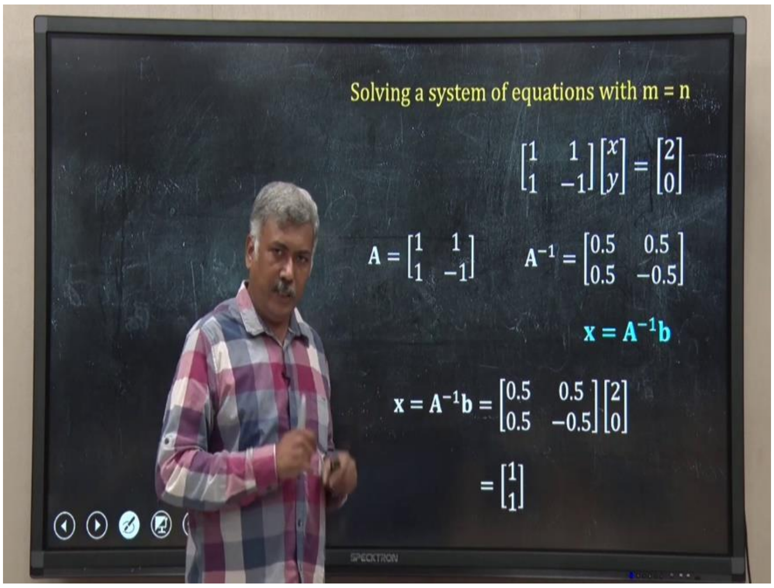
$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}, \text{ iff } \det A \neq 0, \text{ i.e. A must be a full rank matrix}$$

So, what I will do? I will take help of matrix inversion, I will multiply both sides with A inverse. So, what I will get, I will get A inverse into A into X equal to A inverse b, both sides of the equation I have multiply with A inverse. Now, I know A inverse A by definition and if you do any calculation will be in identity matrix. So, what do I get? I get I X equal to A inverse b. Now I into X, identity matrix into a vector will give the same vector because identity matrix is nothing but equivalent to scalar 1.

So, that means I will get X equal to A inverse b. And that is what I want to do. Now, here is a catch. When can you do the inversion of a matrix, if you remember the lecture on determinant and inversion of matrix, I can invert a square matrix only when it is a full rank. That means if its determinant is not equal to 0, and full rank of a matrix means its rows and columns are linearly independent.

Now think back where I am getting the coefficient matrix from? I am getting the coefficient matrix from the system of linear equations. That means to have a full rank coefficient matrix, my system of equations all those equations should be linearly independent to each other. If they are not linearly independent, I cannot do this inversion, determinant will be equal to 0 and I will not be able to get a unique solution.

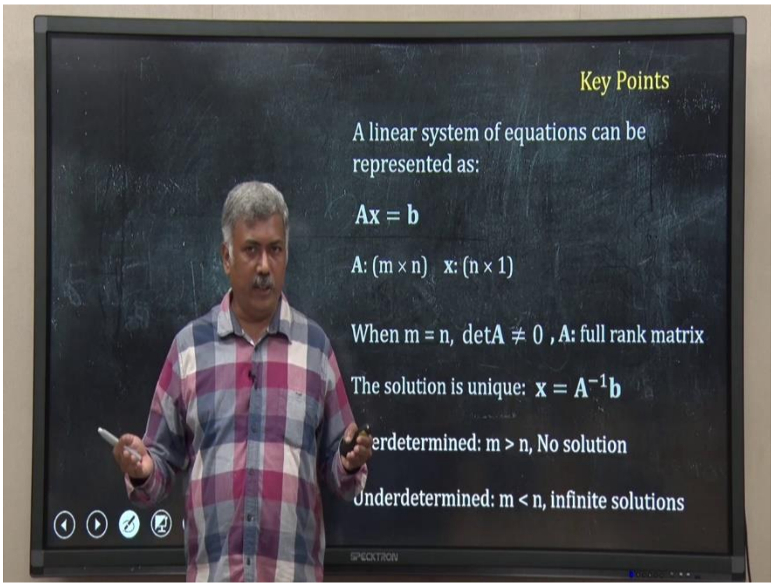
(Refer Slide Time: 25:02)



So, let us first see an example, I have taken a simple system that the first one, I started with x + y = 2, x - y = 0, I have written it in the form of matrix, the coefficient matrix and the vectors. So, now I calculate the determinant, determinant is -2, and that is not equal to 0, that means I can do the inverse of my coefficient matrix. And that inversion I have done using an online tool, you can also use them. So, I got the inversion.

So, now I will multiply A^{-1} with B to get the x vector, that is what I have done $x = A^{-1}b$ and this is a simple matrix multiplication that will give me 1 1. In fact, we have done simple algebra to get that earlier. So, it is matching.

(Refer Slide Time: 25:52)

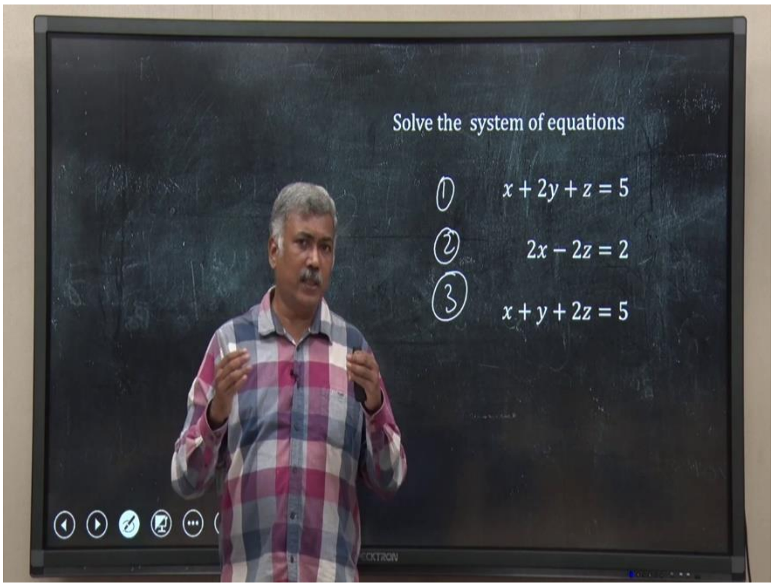


So, now, let me jot down what we have learned in this lecture. In this lecture, we have tried to understand the solution of a linear system of equations. So, first we started with what is a linear system. And then, we have learned how to represent a system of linear equation in terms of linear algebra using matrix and vector as Ax = b, where A is the coefficient matrix of m into n dimension, m is the number of equations, n is the number of variable or unknown, whereas, X is the variable vector having n into 1 size, n is the number of variables.

And then we have learned that if determinant of A is not equal to 0 and A is a square matrix that means, the number of equation is equal to number of variable and as the determinant is not equal to 0, A is a full rank matrix, then a solution unique solution to X is $X = A^{-1} b$.

Along with this, we have also learned about over determined system and underdetermined system. In over determined system number of equation is much more than the number of variables or unknown, whereas in under determined this is the opposite, in overdetermined system, I have no solutions, usually, we encounter them very frequently in data analysis. In that case, we use techniques like regression to find the optimum solution, but that is not a unique solution and underdetermined system m less than n and it is infinite solution is actually not much useful for us.

(Refer Slide Time: 27:44)



So, that is all for this lecture, before I leave, I will give a problem for you to solve, a simple one. So, I have given here one, two and three equations, all are linear. So, you have to solve this system of equations, you can easily use the algebra and replace values and do that like you have done your school math, but I insist that you use the concept of linear algebra that we have learned in this lecture and solve it right.

So, for that you have to calculate determinant, you have to calculate invert of a matrix all these things and they are as I said, there are online tools available, search them out and use them and solve these system of equation using the concept of linear algebra that we have learned in this lecture. I am sure you will try it, then see you in the next lecture. Till then, happy learning.