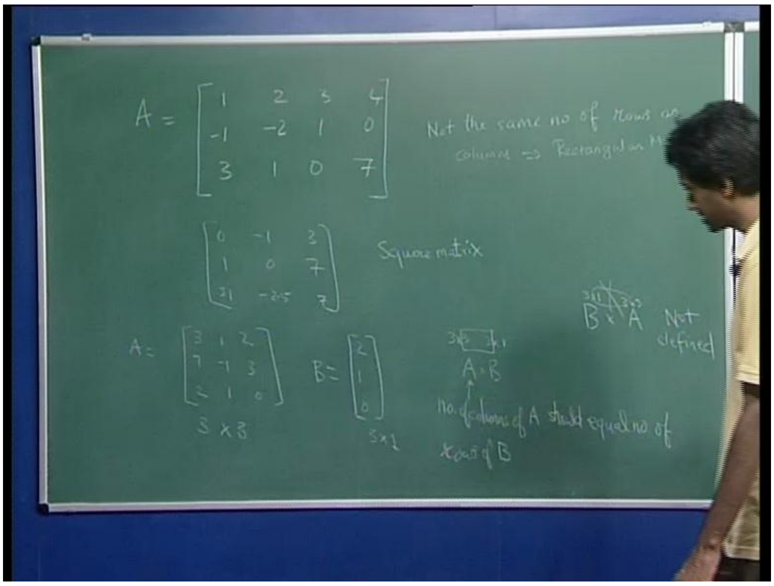
Mathematics for Chemistry

Prof. Dr .M. Ranganathan
Department of Chemistry
Indian Institute of Technology, Kanpur

Lecture - 9

Next, we go to the next topic in this course that is matrices. Now, all of you have seen matrices in some form before. Matrix can be thought of as arrangement of scalars, so it is just a convenient arrangement of scalars. And this arrangement of scalars can be either in the form of a square matrix or it can be in the form of rectangular matrix.

(Refer Slide Time: 00:45)



So, what does it looks like so you must have seen matrices, that look like convenient arrangement of scalar and you can have an example 1 0. So, the scalars, these are the scalars and if you want you can make them real numbers also, and they are arranged in a suitable form and this is called matrix. So, instead of writing all these 12 scalars, just call it a matrix A. Now if you notice, this has a particular arrangement of scalars and these are called rows and these are called the columns.

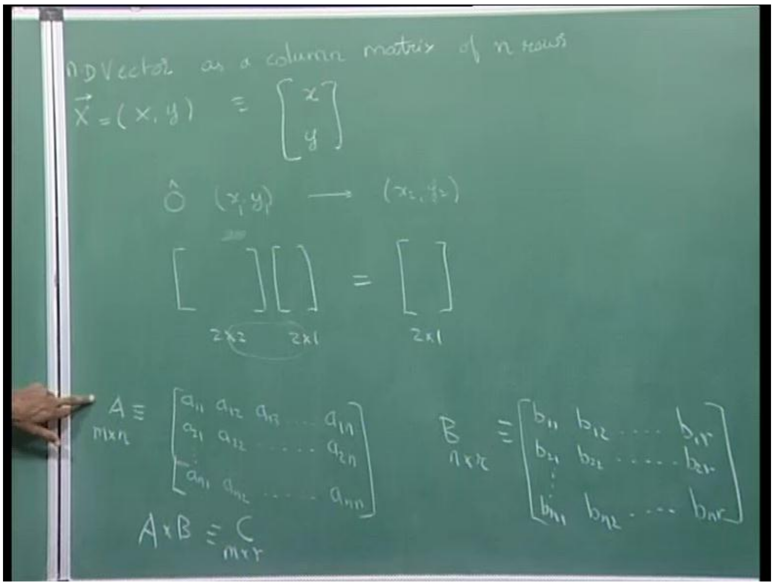
So, in this case, this is a rectangular matrix, so not the same number of rows as columns for rectangular matrix. So, this is the matrix, where the number of rows and columns are not the same on the other hand, you could have a matrix like ((Refer Time: 02:01)) 0, so this is called a square matrix. So, when the number of rows and columns are identical then it is called as square matrix. So, here we have a very visual definition of a matrix, as just a convenient arrangement of scalars, each term here is a scalar alternatively, you can also have it, even be a complex number, so you can have a matrix of complex numbers, you can have matrix of integers, matrix of real numbers and so on. And so far, we have just defined as it is arrangement, now it is also useful to think of a matrix as an operator, that converts a vector into another vector. So, this is another useful way to think of a matrix and in order to that, we need to define how to multiply matrices. So, if you have one matrix A is equal to, so matrix now, this has 3 rows 1, 2, 3 rows and it has 3 columns.

Now, in order to multiply this matrix to another matrix B so it is essential that, the number of rows of B be equal to 3, so b has to be 3 by anything and it can be 3 by let us say, 1 so if it is 3 by 1 so it has 3 rows and 1 column. So, it could be for example, this could be 2 1 0, so it is important that, if I have these two matrices then I can define A into B. So, this is a matrix multiplication of the matrix A and the matrix B, and it is essential that, the number of columns of A should equal number of rows of B.

Now, so in a such a case, you can define A times B but however, you cannot define B times A because the order of matrix multiplication makes a difference. So, in B times A so A is 3 by 3, B is 3 by 1 and you notice that, the number of columns of A is equal to number of rows of B. But, if you try B times A, number of rows of B is so this is 3 by 1 and this is 3 by 3 so this is not defined because these two are not equal.

so, whenever you want to multiply two matrices, you should always check to make sure that, the number of columns of the first matrix is equal to the number of rows of the second matrix. Now, once you have this definition of matrices then there is another way and another useful way to look at matrices.

(Refer Slide Time: 05:48)



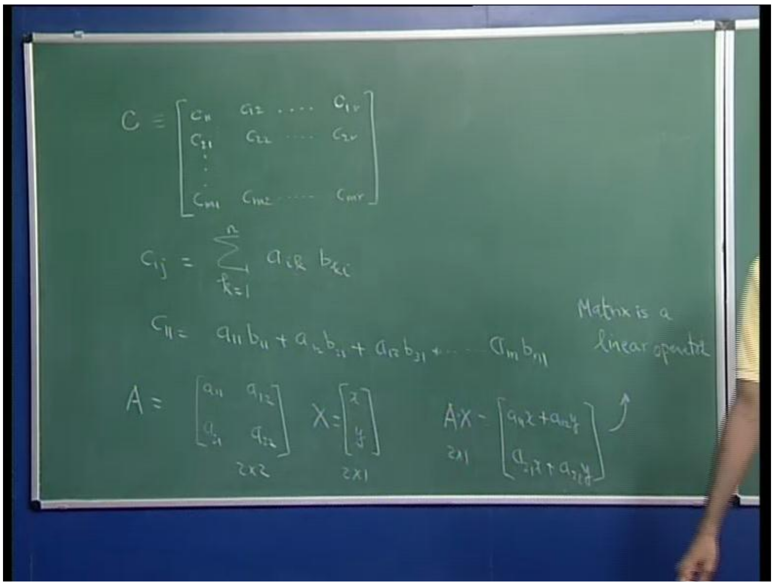
Now, in this way, you can think of a vector as a column matrix so for example, a vector, a n dimensional vector will be a column matrix of n rows. So, for example, if you have a two dimensional vector given by x comma y, you can think of this as column vector x and y. Similarly, you can extend it to 3, 4, 5 dimensions and the usefulness of these is that, you can ask, what is the general operator, that acts on one vector, it acts on vector x 1 and y 1 to give you a vector x 2 y 2.

And one such operator is the matrix so this is a 2 cross 1 so if you think of this as a vector, as a column matrix. So, this is x vector is equal to x y and this the column matrix corresponding to x vector so I can multiply by a 2 by 2 matrix on this 2 by 1. So, if I do 2 by 2 matrix multiplied by the column vector x, which is 2 by 1 then I will get another 2 by 1 matrix. So, if I multiply these two matrices then the order of the resulting resultant matrix will be a 2 by 1 matrix.

And this is chosen so that, the matrix product is defined in the following way suppose, I have a matrix A and a21 of a1m, a2m all the way upto a n. So, an m by n matrix has various elements defined in this way and m by n matrix can multiply another matrix of order n by r. So, if the elements of B are represented as b 11, all the way upto b nr so m by n matrix A can be multiplied to an n by r matrix B and the resultant matrix A times B is an m by r matrix.

So, what we will see is that, A times B gives C where, C is an m times r matrix so let us come back to this, we take a matrix A, which is m by n matrix multiplied with another matrix B, which is n by r matrix to get C, which is an m by r matrix.

(Refer Slide Time: 09:52)



And if you want to find the elements of C so if C has various elements c 11, c 1r, c 21, c m1 so if your matrix C has all these elements then you can write the i j th element of C is equal to sum over a ik b ki, sum over k equal to, and A has n elements so k equal to 1 to n. So, you multiply the one, the i 1 element with the b k 1 so suppose, you were looking at the first element of C so c 11 is equal to a 11 b 11 so where, k equal to 1 plus a 12 b 21 plus a 13 b 31 plus all the way up to a 1n b n1.

So, that is how, we define c 11 so c 11 is essentially this, multiplied by this plus this multiplied by this plus this element multiplied by this and so on, all the way to this element multiplying this element. So, a 11 b 11, a 12 b 21, a 13 b 31, all the way up to a 1n and b n1. Similarly, we can write all other elements now, once you have defined this matrix multiplication in this form then you can ask, what is it mean to multiply a matrix on to a vector.

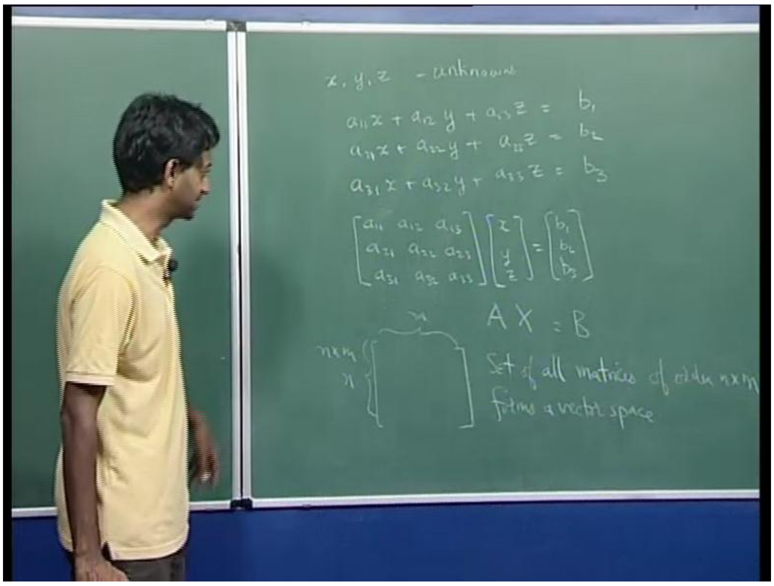
So, if you multiply a matrix with a vector for example, if you have a 2 by 2 matrix, a 12 a 21 a 22, if you take this matrix, if you take a 2 by 2 matrix and multiply it to a vector so if A is equal to this and X equal to x y then AX. So, this is 2 by 2, this is 2 by 1, A into

X, A times X, this should be a 2 by 1 matrix and the elements of the matrix are a 11 x plus a12 y, a 21 x plus a 22 y.

So, what we find is that, if you operate a matrix on a vector, you get another vector so this is 2 by 1 matrix. So, this is also a vector and components of the vector, the first component is a linear combination of these two of the components, of the original vector. Similarly, the second component is also a linear combination of the components of the original vector so therefore, a matrix is a linear operator so this implies that a matrix is a linear operator.

So, if I multiply a matrix on the sum of two vectors, it will be the sum of the matrix multiplying the vectors. So, we have seen two useful ways of looking at matrices, one is in general, as an arrangement of scalars, the other is as an operator acting on a vector to give you another vector. And such an operator is a linear operator because the components of the resulting vector are linear combination of the components of the original vector. Now, one of the most important motivations of matrices and the use of matrix is, in solving a system of linear equations.

(Refer Slide Time: 14:42)

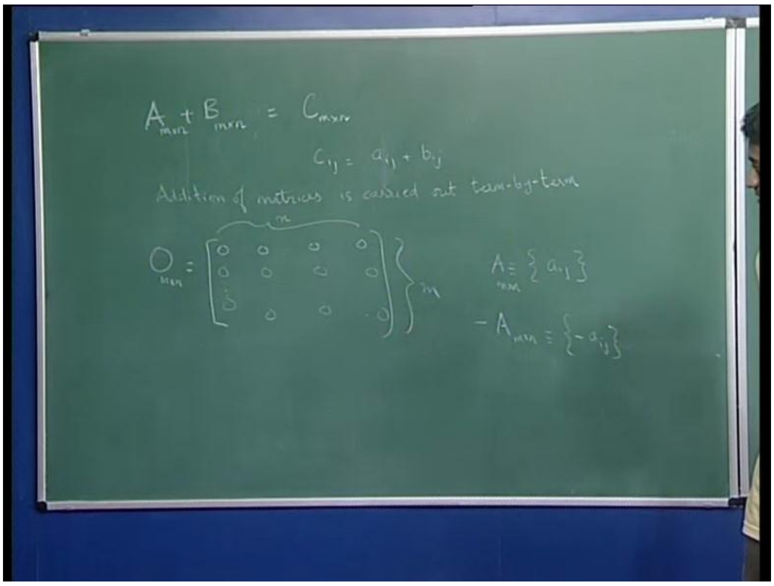


What do you mean by a system of linear equation suppose, you have three unknowns x, y and z are unknowns. Now, a system of linear equations would be various equations involving x, y and z so for example, you can have a11 x plus a12 y plus a13 z is equal to b1, a21 x plus a22 y plus a23 z is equal to b2 and so on. So, you could keep having this for as many as you wants so for convenience, let us take a third one now, this immediately reminds us of matrix multiplication.

And it is not hard to show that, this can be written in this form a31 so you can write it in this form and so here lies utility of matrices. So, system of equations, I can write it in a simply form, if I call this B vector, call this x, I can just write AX equal to B, a matrix A multiplying a vector X to give you another vector B. So, this whole system of linear equations can be written in a very compact form and it turns out that, various properties of this system of linear equations are related to the properties of the matrix, as we will see very soon.

Before, we go into little more details at this system of linear equations, let us mention one thing that, if I take matrix all, if I take a n by m matrix. So, this has n rows and it has m columns, so if I take a matrix like this and if I consider the set of all matrices of order n by m, this set of all matrices forms a vector space. So, this forms a vector space and the reason that it forms a vector space is because addition of matrices is defined by this way.

(Refer Slide Time: 18:08)



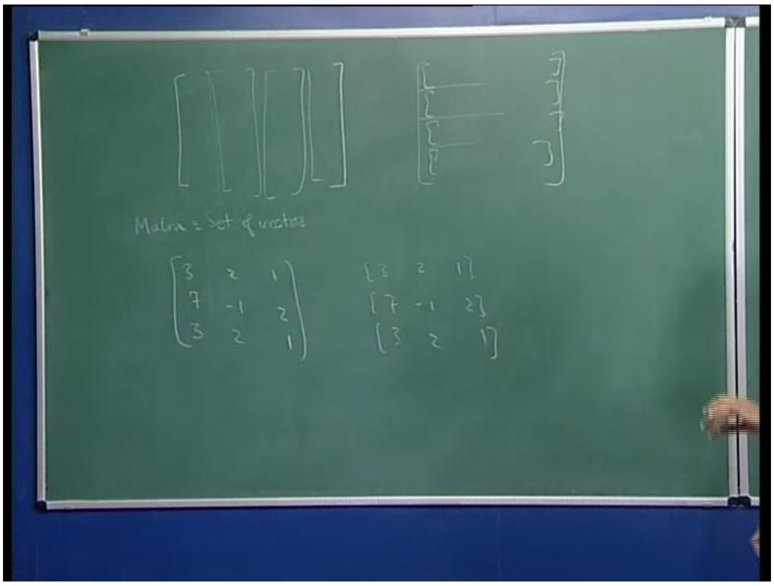
So, you have to define the addition of matrix so addition of matrices, if you have one matrix A plus B and two matrices A and B can only be added, if they have the same dimensions. So, this is m by n, this is m by n then this is the matrix C is also an m by n matrix such that, C ij equal to a ij plus b ij, so the addition of matrices is term by term.

So, the way you add two matrices of order n by m will form a vector space and you can easily show that, this will contain the zero matrix.

The 0 matrix is the matrix with 0, so you have n elements here, m elements here so the 0 of order m by n will just have 0's everywhere. So, matrix with 0's everywhere is the 0 matrix and this is the identity matrix for addition of matrices similarly, the inverse of any matrix is also defined. So, if you have a matrix A is equal to, A m by n is given by terms of the form a ij so A m by n has various elements like a ij. For example, if a 11, a 12, a 13, etcetera then minus A is also an m by n matrix where, the terms are inside minus a ij.

So, you just take negative of each of them so just as you had vectors and you had the negative of a vector, you can have matrix and you can have negative of a matrix and this should be obvious because vector is a especial case of matrix with just one column. So, in this notation, vector is a special case of a matrix with just one column now, next what we are going to see is, one of is a very important application of matrices and in order to do this, we will define something called the rank of a matrix.

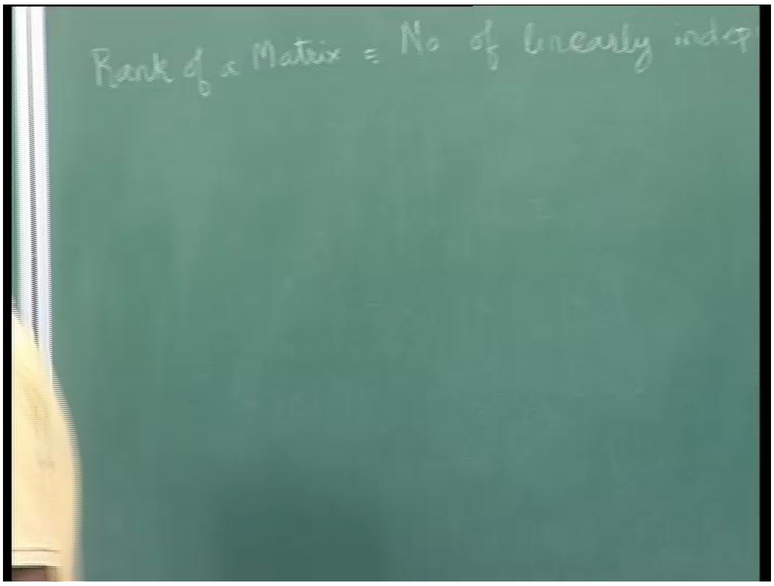
(Refer Slide Time: 21:04)



So, before, we define the rank of a matrix, let us mention one thing in brief, that another way of looking at a matrix is like a, think of a matrix as a series of, as a set of rows, vector. So, matrix equal to set of vectors and these vectors can be either column vectors or they can be row vectors. So, you can think of either each row as a vector or each column as a vector or you can think of a matrix where, each row can be thought of as a vectors.

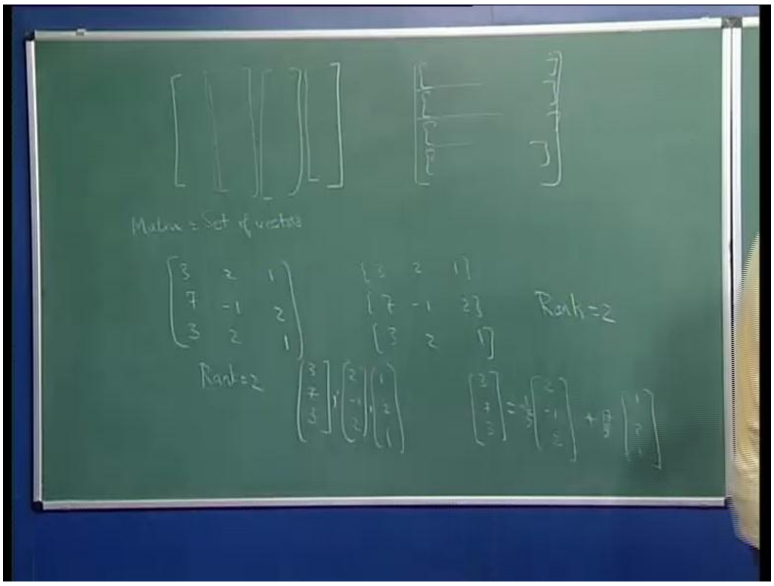
So, the matrix can be thought of either a set of vectors and you can think of either each column as a vector or each row as a vector. So, you can think of this matrix has four vectors so for example, if you have a matrix 3 2 1, you can think of 3 vectors, you can think of it as compose of 3 vectors. And now, that you have these 3 vectors, you can ask questions about are these vectors, you ask questions about the relation between these vectors in particular one question, that will be relevant is, whether these vectors are linearly independent or they are linearly dependent. So, for example, if you take these 3 vectors, are they linearly independent or are they linearly dependent.

(Refer Slide Time: 22:53)



So, and this allows us to define something, called the rank of a matrix equal to number of or rather the number of linearly independent rows or columns. So, the rank of a matrix is, you look at the rows as vectors and you ask, how many of them are linearly independent alternatively, you can look at the columns or column of a matrix and you can ask, how many columns are linearly independent. And it turns out that, it is not have to show that, whether look at row or column, you will get the same answer and this answer is called the rank of a matrix.

(Refer Slide Time: 23:54)



So, for example, in this case, for this matrix you see that, these two rows are identical, since these two rows are identical, these two vectors are not linearly independent. However, these two vectors turn out to be, it is not have to show that, these two are linearly independent so the rank equal to 2 so the rank of these matrix equal to 2 so here I should say, rank equal to 2.

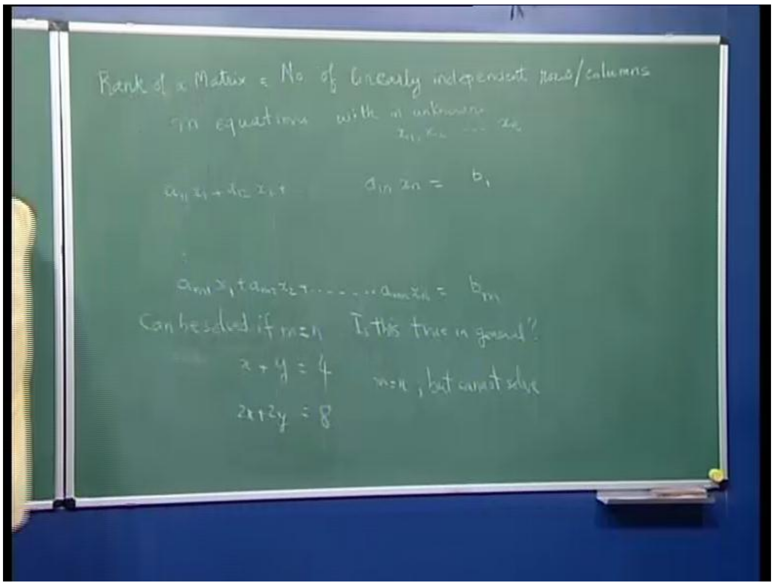
So, now, if you had instead, chosen to work with column vectors so if you have chosen to do it by column vectors, you would say 3 7 3 2 minus 1 2 1 2 1 and it should not be hard, if I multiply. So, you take 1 2 1, 2 minus 1 2 and 3 7 3 and it should not be hard to show that, these three are actually linearly independent. So, you should be able to construct linear combinations of these, you should be able to express one as a linear combinations of the other.

So, for example, you should be able to write 3 7 3 is equal to some constant into 2 minus 1 2 plus some other constant into 1 2 1 and let us see, if I take this as 5 then this should be ((Refer Time: 25:50)). So, you should be able to determine this constants by choosing suitable linear combinations of these and this exists for some so you can determine constant c and c prime such that, this equality is satisfy so the rank of a matrix is the number of linearly independent rows or columns.

So, the point is now, by suitably choosing values of c prime and c, you can get this equality satisfied and it is not hard to show that, what should be the value of c prime should be, should be nothing but 17 by 5. And then this can be c, c can been shown to be minus 1 by 5 so then this is minus 2 by 5 plus 17 by 5 that is, 3. Similarly, this is also minus 2 by 5 plus 17 by 5 that is, 3 and this will work out to 1 by 5 plus 34 that is, 35 by 5, this is 7.

So then you can show that, these three column vectors are not linearly independent they are in fact, dependent and you can also show that the rank of the matrix is equal to 2. So, the rank of the matrix is the number of linearly independent rows or columns.

(Refer Slide Time: 27:26)



And suppose, I consider a matrix where, the number of rows and columns are different so suppose, I consider 4 by 2 matrix. So, if I consider 2 minus 1 3 0 4 minus 1 7 1 now, the number of linearly independent columns is 2 so if you look at columns then you cannot have more than two and in fact, these two columns are in fact, linearly independent.

What about the number of rows, how many rows should be linearly independent and it is not hard to show that, since each row is just two dimensional vector, the maximum number of linearly independent two dimensional vectors is 2. So and in fact, it turns out that, there are only two linearly independent vectors in this side so whether you look at rows or columns, the rank of the matrix is identical. So, the number of linearly independent rows and the number of linearly independent columns of a matrix are the same.

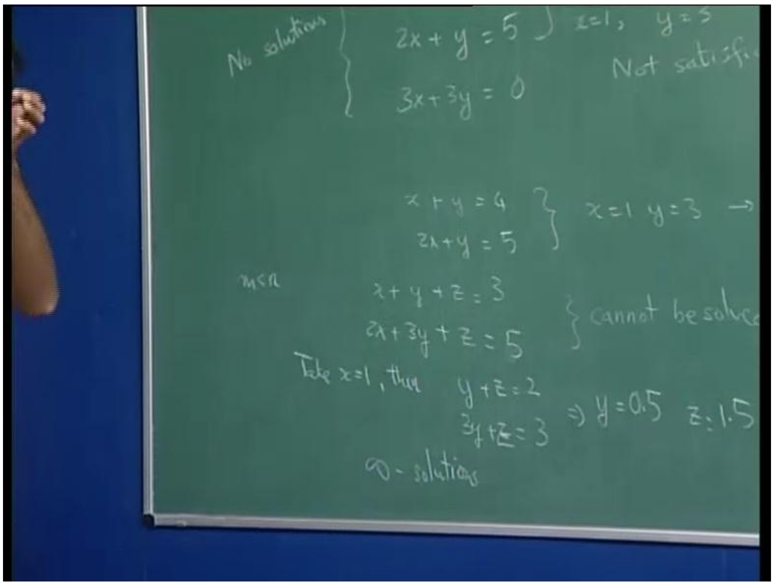
So, how do you use the idea of rank of a matrix and the most important application of the of the use of rank of a matrix is an solving a system of linear equations. So, let us say, you have a system of m equations with n unknown now, we will call the unknowns, will call the n unknowns as x 1, x 2 up to xn. Now, you have m equation involving this n unknowns so the most general linear equation will be a 11 x 1 1 n x n is equal to b 1 and you have a set of m such equations.

So, you have n equations and sorry m equations and n unknowns, and the question is, you are asked to solve and find the values of all the variables x 1 to x n. Now, the first question you will ask is, can you solve these set of equations and if you can solve it then the next step is to actually go ahead and solve it. So, let us look at some specific examples and then we can look at the details so the first thing you will say is that, since I have n unknowns, if I want to solve it typically, I should have n equations.

If I had n equations and n unknowns then I would be able to solve it so you would say, it can be solved if m equal to n and this is the first guess that, you will say. So, now, let us look at some examples, if m is greater than n then we say that, I have too many equations and too few unknowns so it may or may not be solve. In some sense, I have too many condition and too few unknown so actually the system is say, over determined whereas, if m is less than n you say that, I have too few equations so I cannot solve for all the variables.

But, let us look at this in a little bit more detail with specific examples so let us look at a case where, m is greater than n or first, we will start with the case where, m is equal to n. So, the question is, is this true so suppose, I take the equation x plus y equal to 4 and I will say, 2 x plus 2 y equal to 8. So, I have two equations, I have two unknowns but if you try to solve it, you can immediately see that, you cannot solve this set of equations. The reason is that, if I divide by 2 here, I get back this same equation, as I started with so m is equal to n but cannot solve.

(Refer Slide Time: 33:08)



So, next let us consider another system of equations let us say, x plus y equal to 4, 2 x plus y equal to 5 and we have a third equation 0. So, we have these now, we have three equations and you have only two unknowns so in general, you would say that, this is an over determined problem. So, what happens in this case, let us say, you take the first two equations and when you look at these two, you will get x equal to 1, y equal to 3.

Now, when you substitute this in the third equation, you do not get an answer so not satisfied in other words, what you say is that, the third equation is not consistent with the first two equations. So, you say that, these three set of equations are not consistent so let us come back to these two equations now, these two equations I said that, you cannot solve them and in this case, we say that, the third equation is not consistent with the first.

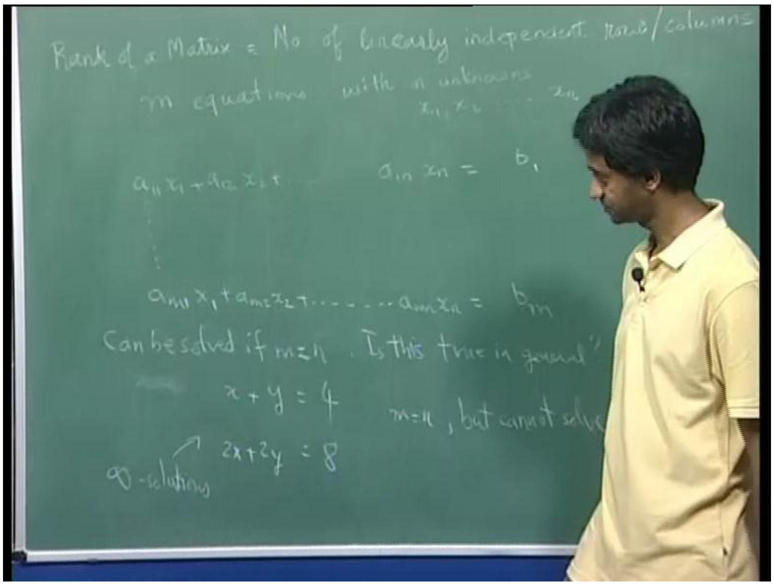
Now, when you say, you cannot solve for them suppose, I take x equal to 10 then y equal to minus 6 will satisfy both the equations. So, for any value of x, I can find a value of y so in other words, there are infinitely many solutions, infinite solutions in this case. Now, in this case, you can never find so this has no solutions on the other hand, if you just add the first two equations, this implies this has unique solution, this is unique solutions.

Now, so we seem to have a set of equations where, m is either equal to n or m is greater than n but in each case, you have different characteristics. In one case you have infinitely many solutions, in the second case you have no solution and in the third case you have a unique solution. Now, if m is less than n for example, if you have x plus y plus z equal to 3, 2 x plus 3 y plus z equal to 5 and let us make it 5.

So, here, you have three unknowns and two equations and you say that, you cannot solve this in other words, if I pick one value of z then I can always find a set of x and y. So, again this cannot be solved but suppose, I take x equal 1 then I have y plus z equal to 2 and 3 y plus z equal to 3 and these two equations can actually be solved. So, these two equations can be solved and you can get solution so for example, take x equal to 1 then you get y plus z equal to 2 and 3 y plus z equal to 5.

So, in other words, you get 3 y plus z equal to 3 so then y plus z equal to 2, 3 y plus z equal to 3 so you can conclude that, y equal to 1.5 or equal to 0.5, z equal to 1.5 so this is a solution. So, this solution holds, when you take x equal to 1 similarly, you can take any value of x and you can solve this and so in other words, you will say that, there are infinitely many solutions.

(Refer Slide Time: 38:31)

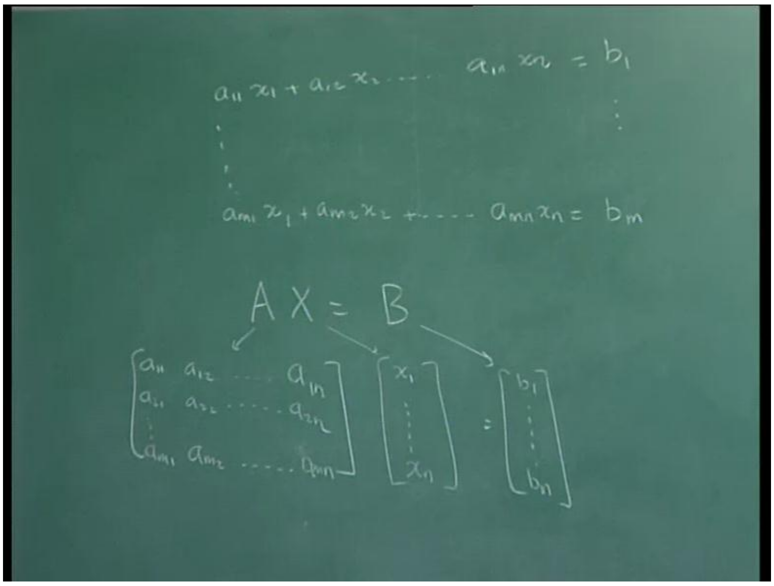


So, basically, it seems from all these argument that, your original idea, that the number of equations should be equal to the number of unknown is not a complete idea. Because, in some cases, you can solve it, if the number of equations and unknowns are the same, you can solve it in some case, you cannot solve it, in some case there are infinitely many solutions and they seem to be too many cases.

And this simple criteria that, number of equations and unknowns should be the same, does not seem to hold in general. And so this is where, the rank of the matrix comes into play and what we notice is that, when you said that, these two equations they cannot be solved, even though number of equations and unknowns are the same, we notice that, these two equation are completely linearly dependent.

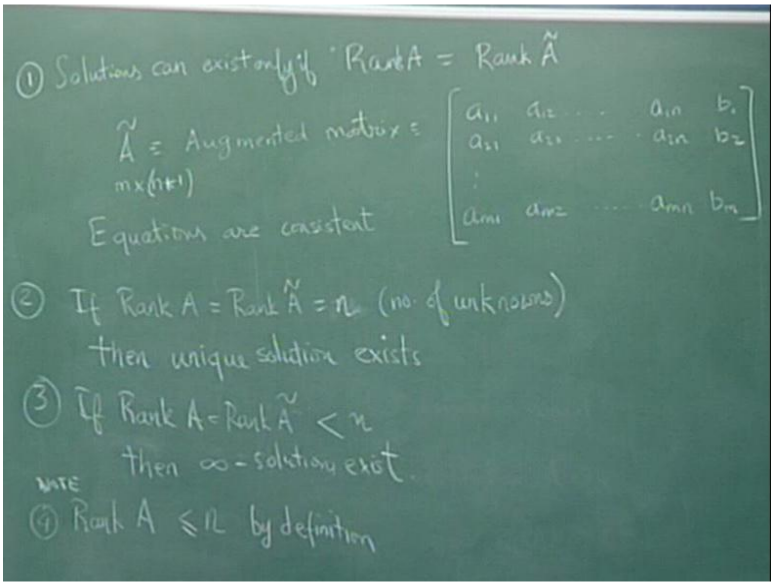
So, this equation is just twice this equation and in a sense, you have only one independent equation and in this case, what happened was, this equation was not consistent with these two. So, that is another case that, we have to consider and finally, in this case, these two equations could be solved because you have two equations and two unknowns and so on. Let us see, how to state this in very simple way now, I am writing my system of equations.

(Refer Slide Time: 39:59)



So, if I have a 11 x 1 plus a 12 x 2 a 1n x n and if I go all way to a m1 x1 a m2 x2 a mn x n equal to b m, this equal to b 1. So, if I have a system of equations of this form then I can write this in the form A X equal to B where, A is this matrix of coefficience. So, A is a 11 a 21, all the way up to a m1 a m2 a m n. So, this is my A matrix, X is x 1 up to x n and B is b 1 up to b n now, what we can say about this system and the solvability.

(Refer Slide Time: 41:33)



So, the first thing will say is that, solution exists if rank of A is equal to rank of A tilde where, A tilde is called the augmented matrix and this has coefficient so it looks like a 11 a 12 a 1 n and then you put b 1. So, what you do to construct this matrix is, if you take this vector and put it here so you have a m cross n plus 1 so there are n plus columns to this.

So, this had m cross n, this is n cross 1, n cross 1 so what you did is, you took this whole b b matrix and you attached to the side of a and you created this larger matrix a 2n b 2 all the way upto. So, now, the first thing you say is that, the solution exists if rank of A equal to rank of A tilde so only in this case, can the solutions exist. So, if this is not true then it means that, this equation are not consistent with each other so if this is not true then the equations are not consistent with each other.

So, if this is true, your equations are consistent so if this is true then the equation are consistent and this is a necessary condition for any solutions to exist. So, probably, I should write this as so right way to say this is, solutions can exist, only rank A equal to rank A tilde. So, the first thing is here that, whatever equations I give should be consistent with each other and that is true, if only rank of A is equal to rank of A tilde.

Now, if rank A equal to rank A tilde equal to m, if rank A is equal to rank A tilde, is equal to the number of unknowns, this is the number of unknowns sorry this should be yes or it should be, n number of unknowns then unique solution exists. And the third thing is, if rank A is equal to rank A tilde and this is less than n then infinitely many solutions exists.

So, the rank tells you the number of linearly independent rows or columns so if the rank A is less than n that means, you have less than n independent equations so that means, you can have infinitely many solutions. So, in a very simple way, using the concept of rank, you can find out, in which case are the equations consistent or which is, what is necessary for solution to exist and you can find out, whether you have unique solution or you have infinitely many solutions.

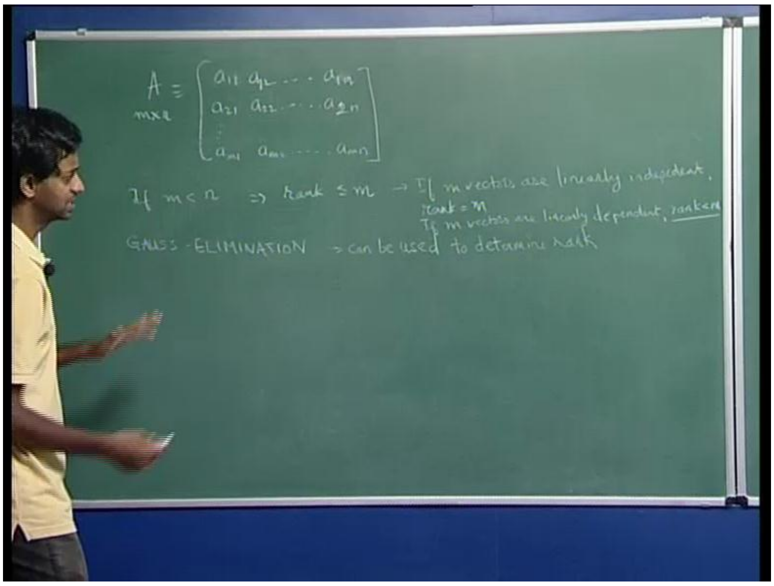
And it is not hard to show that, since the number of rows of A is equal to n, the rank can never be greater than n. So, rank of A can never be greater than n so you can never have rank of A greater than n so we will just put a note saying that, rank of A has to be less than equal to n so this is by definition so this is a note. So, you only need to consider, what happens if rank A is less than n and what happens, if rank A is equal to n.

And the condition for unique solution is that, rank A is equal rank A tilde equal to n and otherwise, you can have infinitely many solutions. So, the rank really tells you about the solvability of a system of linear equations and this is actually one of the most used concept in various problems, in chemistry. We finally, end up formulating your problem as a set of linear equations and then you look at the solvability based on the rank.

So, we have talked about the rank and we have said, how the rank can be used to tell, whether a system of equations has a solution or not. Now, two questions remain, one is the rank tells us, whether you have solutions or not now, what is the best way to determine the rank of a matrix. We have said that, we have a set of vectors and you have to determine how many of them are linearly independent and it turns out that, there are some good ways to determine the rank of a matrix.

Second thing that remains is that, suppose, there exists a unique solution then what is the unique solution so what is a solution. So, these are the two question that, we will consider next so let us start with determining the rank of a matrix.

(Refer Slide Time: 48:41)



So, if you had given A a 21 a 1n this is 2n a m1 a m2 up to a mn so you have a matrix and m by n matrix and you are asked to determine it is rank. So, the first thing you see is reaches so you have to find out, how many rows or columns are linearly independent. Now, if m is less than n that means, the number of rows is smaller than the number of columns that implies, rank less than equal to m so then what you can do is, you can take each of these rows and you can find out, how of them are linearly independent.

So, you can take all these rows and find out, how of them are linearly independent now, this in general, it can be a very tedious operation. You essentially take linear combinations of these rows and set it equal to 0 and then you go through the usual process of determining linear independence. Fortunately there is a more elegant way to do this and this is by a procedure called Gauss elimination and the reason for using this, we will discuss this in the next lecture.

But, the reason for using Gauss elimination is that, suppose, you take all these vectors and you look for linear independence. If they are all linearly independent so if rank is less than m, if m vectors are linearly independent, rank equal to m however, if they are linearly dependent, if are linearly dependent rank less than m. So, all you can say is, the rank is less than m but you do not know what is the value of this rank?

So then you have to consider various subsets of vectors and see, how many of them, what is the maximum number of them, that are linearly independent. So, you go through this procedure and if you have large matrices, you can see that, this procedure can get a little tedious. So, there is more elegant procedure called Gauss elimination, which we will discuss in the next lecture.