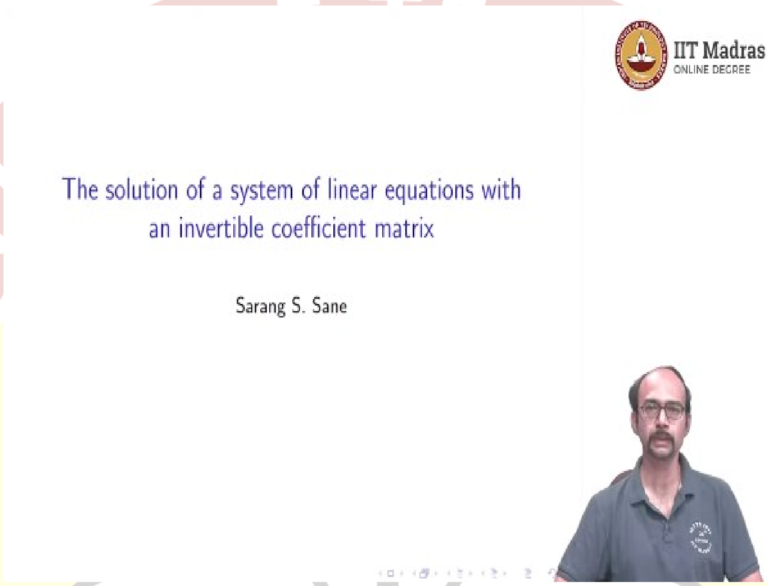
INDIAN INSTITUTE OF TECHNOLOGY MADRAS
सिद्धिर्भवति कर्मजा

IIT Madras

ONLINE DEGREE

Mathematics for Data Science 2
Professor Sarang Sane
Department of Mathematics
Indian Institute of Technology Madras
Lecture 08
The solution to a system of linear equations with an invertible coefficient matrix

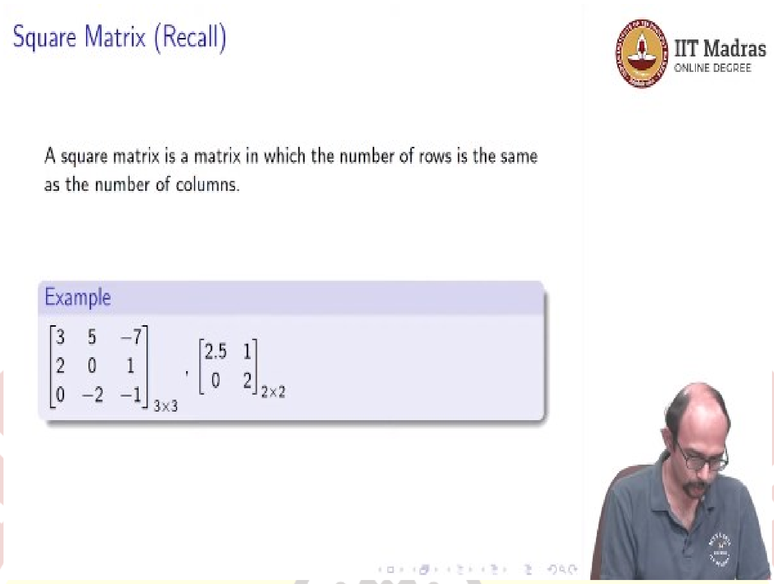
(Refer Slide Time: 0:15)



Hello and welcome to the Maths 2 component of the online BSC degree in Data Science. In the last video we have seen something called Cramer's rule which was used to solve a system of equations in which the coefficient matrix is square and its determinant is non-zero.

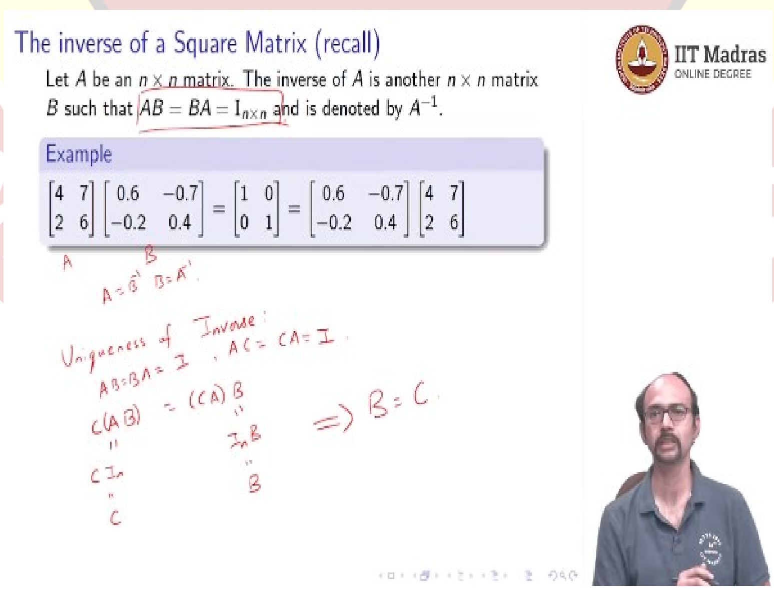
Today we are going to continue on that theme and we are going to talk about the solution of a system of linear equations with an invertible coefficient matrix. So, we hinted on this in the last video and we want to continue with that theme. So, for quite some time we will talk first about the invertibility of a square matrix, what is its relation with the determinant and then move on to the answer to this question about what is the solution and how we can get it.

(Refer Slide Time: 1:07)



So, let us recall quickly what is a square matrix? A square matrix is a matrix in which the number of rows is the same as the number of columns. So, here is an example, so the first one is a 3 × 3 example, the second one is a 2 × 2 example. So, a non-square matrix example would be something which has 2 rows and 3 columns or let us say 5 rows and 8 columns.

(Refer Slide Time: 1:32)



So, let us recall what is the inverse of a square matrix. So, if you have an n × n matrix then the inverse of A is another n × n matrix B such that A×B = B×A = I_{nxn}. Now, I will point out that B has to be n x n otherwise this multiplication on both sides is not defined as, the number of columns of the first matrix has to be the same as number of rows of the second matrix in order for them to be multiplied. So, that forces B to have the same number of rows and columns as A, if this identity were to be satisfied and recall of course, that the identity matrix is the diagonal matrix with all ones.

So, the inverse is denoted by A^{-1}, now I am using already in this. The inverse of a square matrix, so I should qualify 'The' before that, let us see an example, so here is an example, so you have the matrix $\begin{bmatrix} 4 & 7 \\ 2 & 6 \end{bmatrix}$ and you can take B to be $\begin{bmatrix} 0.6 & -0.7 \\ -0.2 & 0.4 \end{bmatrix}$ of course, you can wonder where did I come with these values from, was it trial and error or how did I come up with them. So, that we will study that in a few slides.

Anyhow if you multiply these you get identity and if you multiply it on the other side also you get identity, so this second matrix is the inverse of the first matrix and similarly, the first matrix is the inverse of the second matrix. So, if I call this A and this B then $A= B^{-1}$ and $B = A^{-1}$.

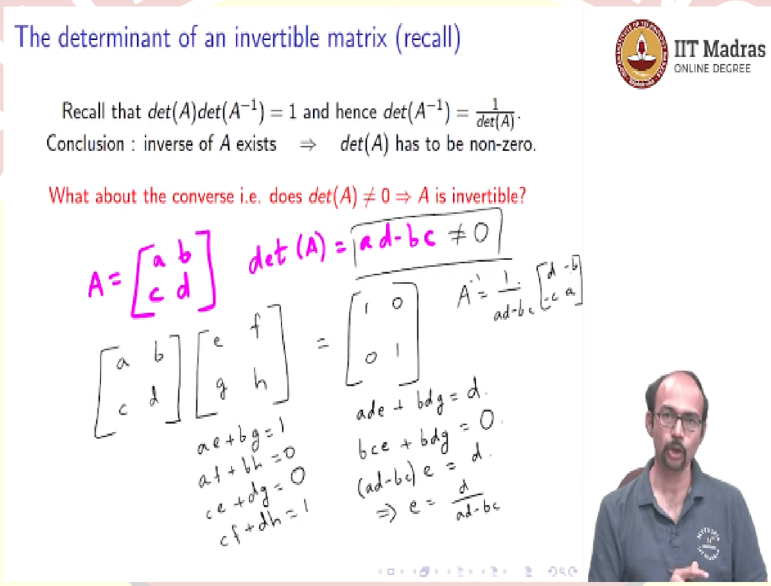
Now, how do I know that there is no other matrix which satisfies the same thing? So, maybe A^{-1} B has this and then some other C also could have the same property. So, how do I know that A^{-1} is this matrix and not some other matrix C? So, this is what is called uniqueness, so the inverse is unique which is why I am using the word 'The' and this is a standard argument in something called group theory but let us go through it now.

So, suppose there are two matrices with this property so AB = BA is the identity and AC = CA is the identity, then what do we do? We compute this matrix CAB, first of all note that all of them are n x n matrices, so this product makes sense, second one of the properties that we saw for matrix multiplication was that it does not matter which order you, it was that, it does not matter whether you first do AB or first do CA and then multiply that to B.

So, in other words we could either do this or we could do this and they are both the same, but let us compute each of these and see what we get. so, if you first do A times B, then you get well A times B is the identity matrix, so C times the identity matrix but anything times identity is itself. So, C times identity is C, on the other hand if you first compute C times A you get the identity matrix times B, but identity times any matrix is itself, so we get B.

So, what did this show? This showed that B must be same as C. So, in other words if two matrices satisfy this relation that AB = BA is identity, then they must be the same, so we are saying that the inverse is unique. Of course, we do not know if it exists or not, as we saw there was a condition for it to exist.

(Refer Slider Time: 5:27)



Namely, we saw that the determinant must be non-zero, how did that come about? That was because we saw that the determinant is multiplicative, meaning det(AB) = det(BA). So, if you take det(AA^{-1}) that is det(I) which is 1, so that means in particular that neither of them can be 0, they have to be non-zero otherwise the product cannot be 1, 0 times anything is 0 as we know.

So, this is a necessary condition for a matrix to be invertible. So, if the determinant is 0, so determinant of A is 0, then the matrix A cannot be invertible, that means it cannot have an inverse. So, we say a matrix is invertible if it has an inverse. So, the conclusion here is that the inverse of A exists implies that the determinant of A has to be non-zero, we can go the other way and ask what about the converse? So, if the determinant is non-zero is A invertible, meaning does an inverse of A exist? So, let us maybe do an example.

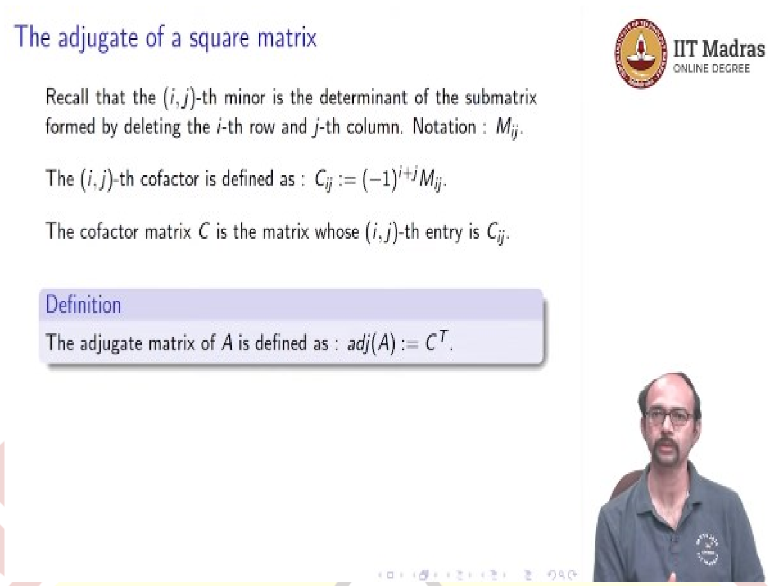
So, let us do a 2×2 example. So, suppose A=$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and I am trying to look for its inverse and what do I know? I know determinant of A is non-zero, so determinant of A is well let us calculate what that is, that is ad-bc, that is how we define the determinant of a 2×2 matrix. So, now I know that this is non-zero, that is what I am given.

So, now let us try and find an inverse. What does the inverse satisfy? So, the inverse satisfies that if you take a, b, c, d and if I multiply suppose, there is an inverse and it is called $\begin{bmatrix} e & f \\ g & h \end{bmatrix}$ then I should get $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. So, that means I will get a bunch of equations and I can try to solve them. So, what equations do I get? I get ae + bg = 1, af + bh = 0 and then ce + dg = 1 and cf + dg = 0, and not 1 and cf + dh = 1. So, this is what we get, these four equations are what we get from the left hand side, that is how we do product of matrices.

Maybe I will multiply the first equation by d. So, ade + bdg = d and maybe I will multiply the third equation by b, so bce + bdg = 0, that is what I get. So, now I can subtract out the bdg, so if I do that what do I get? I get (ad – bc)e = d, so what does that tell me? That tells me that e = $\frac{d}{(ad-bc)}$ and this is the important part, I can divide because (ad – bc) is non-zero, that is the assumption otherwise I would not have been able to solve this at all.

And then you can work out the rest, so from here what we will get is that the inverse of A has this form, so it is $\frac{1}{(ad-bc)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ and I hope my signs are correct. So, my claim is A inverse is going to be of this form and I will leave you to check that the other entries give you exactly this. So, you can see that for a 2×2 matrix if the determinant is non-zero, then indeed the matrix is invertible, it has an inverse, so can I go beyond? So, the converse is true, at least for 2×2 matrices, so what we are going to see next is how to go beyond the 2×2 case.

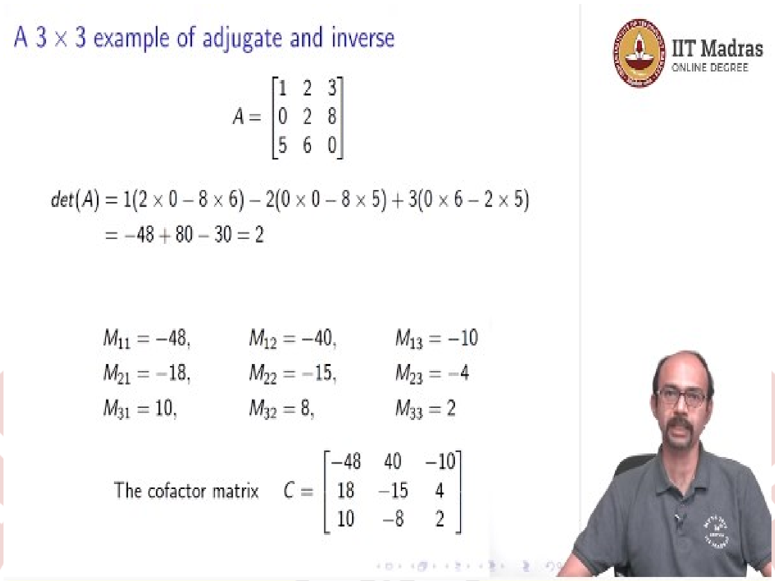
(Refer Slide Time: 10:51)



So, I am going to introduce a few notations, so the adjugate or adjoint of a square matrix, so what is that? So, remember in the previous video, we defined something called minors and cofactors, so the $(i, j)^{th}$ minor is the determinant of the sub matrix obtained by deleting the i^{th} row and the j^{th} column and we denoted it by an M_{ij}. And the $(i, j)^{th}$ cofactor was defined as multiplying the minor, the $(i, j)^{th}$ minor by $(-1)^{i+j}$.

So, we can create a matrix called the cofactor matrix, so the cofactor matrix has $(i, j)^{th}$ entry C_{ij} and then we define the adjugate of this matrix A or adj(A) that you started with as the transpose of this cofactor matrix. So, remember that we defined transpose, so transpose was when you take a matrix and you flip the entries about its diagonals, you reflect the entries about the diagonal.

(Refer Slide Time: 12:04)



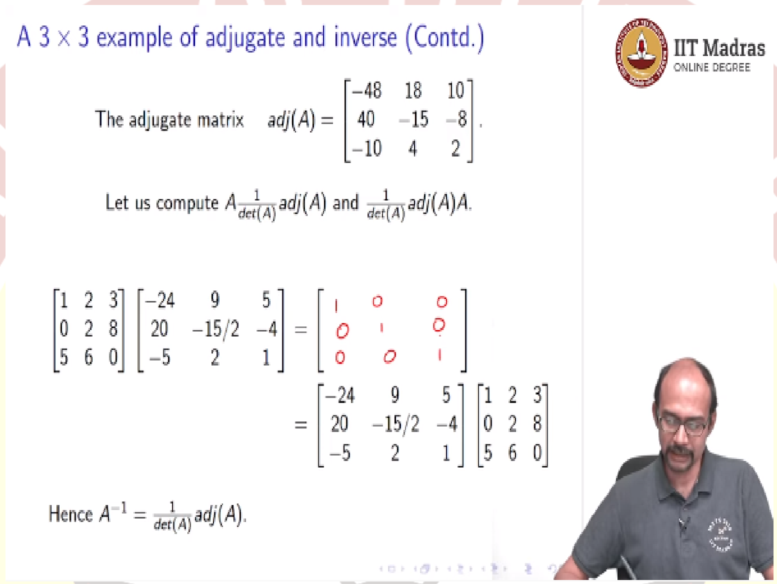
So, let us see an example of the computation of the adjugate matrix. So, here is the matrix $A= \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 8 \\ 5 & 6 & 0 \end{bmatrix}$. Let us compute its minors first, let us compute its determinant first, determinant of A is 2.

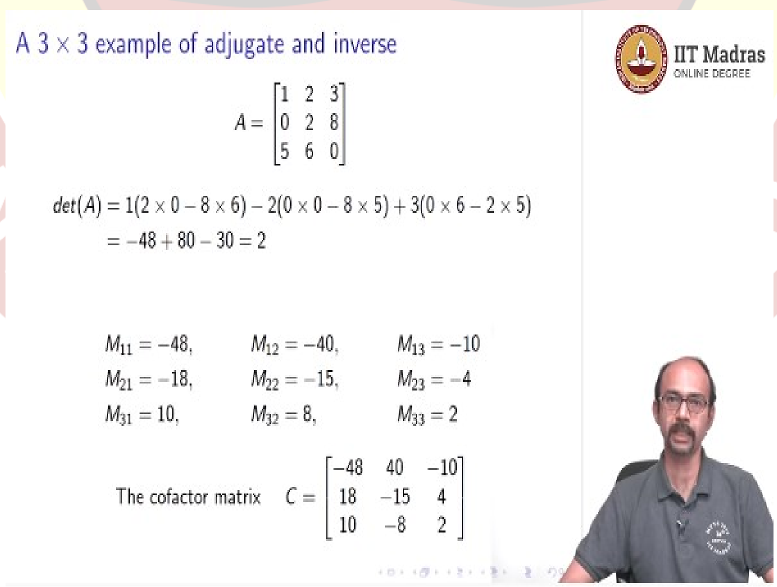
So, let us now work out the minors, so well the minors are written here, I guess you have to work out these minors by yourself. So, just to check that the M_{11} is correct, that is exactly the number appearing in the bracket corresponding to 1 in that expression for the determinant of A. so, that is -48, 1 times -48, the first term in that expression for determinant of that is -1, that is exactly M_{11}, why is that M_{11}? Because that is what you get by hiding the first row and the first column.

So, you can check the other entries, so the other minors and hopefully you will get these numbers. So, now let us write down the cofactor matrix, so the cofactor matrix remember is given by taking these minors and putting them into the corresponding positions but by multiplying each of them by $(-1)^{i+j}$. so, the first term remains the same because it is $(-1)^{1+1}$, so it is -48, the second term comes with a minus sign so its –(-40) which is 40, third term is again 1 times -10 so just -10 and so on.

So, you have 18, -15, 4 and then 10, -8, 2. So, this is your cofactor matrix, again this is something you have to compute by yourself but I hope the ideas are clear, that to compute the adjugate you have to compute the minors first, to compute the minors for M_{ij} you hide the $(i, j)^{th}$ row in the j^{th} column and compute the resulting matrix and then multiply each by $(-1)^{i+j}$ and put that into a matrix that will give you the cofactor matrix.

(Refer Slide Time: 14:36)





And then to get the adjugate matrix you take this same cofactor matrix, let us just see what that was, that was minus $C = \begin{bmatrix} -48 & 40 & -10 \\ 18 & -15 & 4 \\ 10 & -8 & 2 \end{bmatrix}$ and you flip it over the diagonal, you take its transpose. So, your -48, 40, 10, minus 10 became now the first column. So, just to check here, it is the first row, here it is the first column and so on. So, your second row became the second column, your third row became the third column, so 18, -15, 4 here is your second row that becomes a second column and then 10, -8, 2 becomes a third column.

So, you have taken the transpose of the cofactor matrix. So, for your matrix A we have computed the adjugate matrix. So, now let us do something funny, let us compute A$\frac{1}{det(A)} adj(A)$. So, this $\frac{1}{det(A)} adj(A)$ is supposed to be a matrix and that matrix multiplied by A.

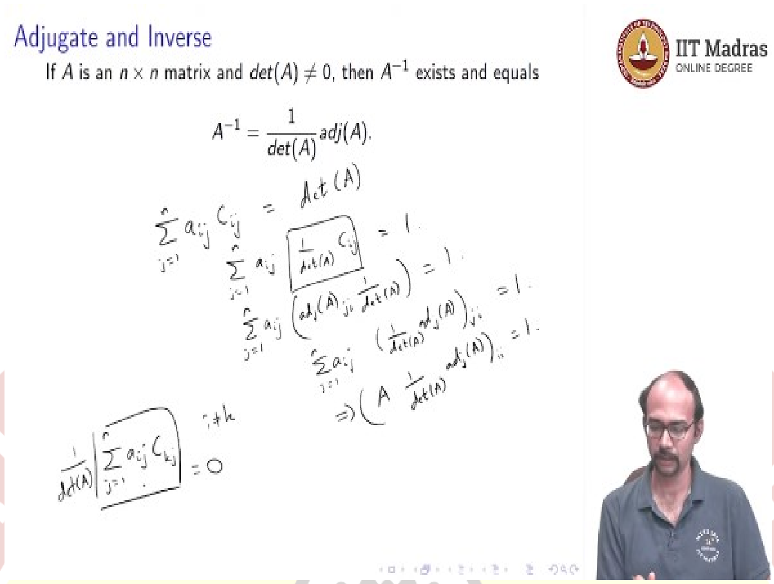
So, let us do this computation, well$\frac{1}{det(A)} adj(A)$, so we computed what is the determinant of A, that was 2, so that means you have to multiply adjugate by half, so if you multiply the adjugate by half the first row becomes -24, 9, 5, the second row becomes 20, -15 by 2, -4, the third row becomes -5, 2, 1 and we want to multiply on the left this same matrix by 2, by A.

So, your matrix A was something you can check from your previous slide. So, now let us multiply and see what we get, so let us let me do the 1×1 term and then the rest I leave to you to check what we or some of the terms and the rest I leave to you to check what we get. So, the first term is 1(-24) + 220 + 3(-5) = 1.

And then the next term, let us say the, so 1×9 + 2× (-15/2) + 3×2 = 0 and you can carry on and do the other terms so you are going to get an identity matrix. So, this is going to give you the identity matrix and now you can do it the other way and you can check that you still get the identity matrix.

So, if you multiply $\frac{1}{det(A)} adj(A)$and multiply A on the right you are still going to get the identity matrix. So, what does that tell you? That tells you that the inverse, A^{-1} is for this example, $\frac{1}{det(A)} adj(A)$, well it turns out that this is a general fact and that is exactly what we are going to state here.

(Refer Slide Time: 18:14)



So, if A is an n x n matrix and determinant of A is non-zero, then A^{-1} exists and equals $\frac{1}{det(A)} adj(A)$. So, how do we get this relationship? So, if you remember in the previous video, we had these 2 very nice formulae about cofactors. So, one was that $\sum a_{ij} C_{ij}$ so this is determinant of A where you take the sum from 1 through n. So, now let us divide both sides by determinant of A, if you do that and then we will take the determinant of A inside the sum.

So, if you do that what we get is $\sum_{j=1}^{n} a_{ij} \frac{1}{det(A)} C_{ij}$ =1 and now we can write this term over here as the ji-th term of the adjugate, so this is the same as $adj(A)_{ji} \frac{1}{det(A)}$, why is that? Because C_{ij} is the same as $adj(A)_{ij}$ because adj(A) is the cofactor matrix transpose.

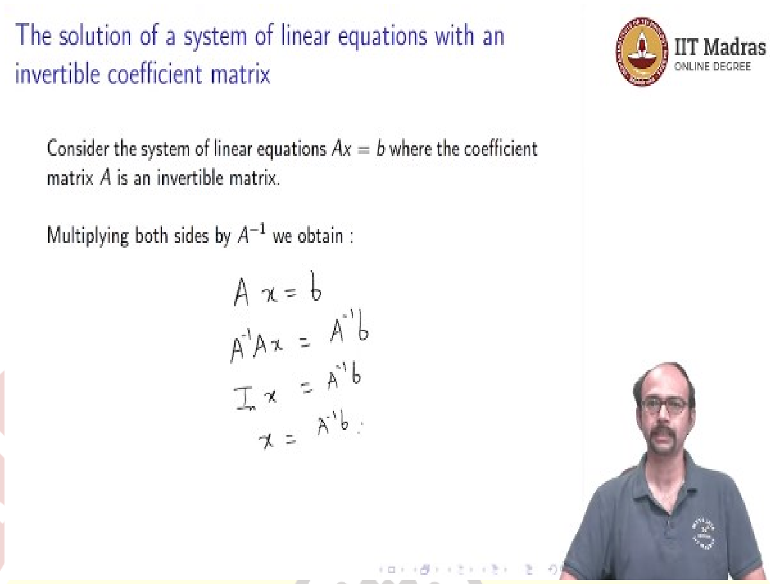
So, this is 1 but this is exactly the, $\sum_{j=1}^{n} adj(A)_{ji} \frac{1}{det(A)}$ =1, so what does this mean? This means that if you take the ii-th entry of the matrix A$\frac{1}{det(A)} adj(A)$ is 1, the diagonal entry is 1. So now we can ask what happens to entries which are not on the diagonal. So, for that we will have to do $\sum_{j=1}^{n} a_{ij} C_{kj}$, why do we care about this, because if you work out in the same way, if you back calculate you will get here.

So, I will leave it to you to back calculate, so what I am saying is you take this matrix and look at its ik-th entry, where i is not equal to k, then this is exactly what you are going to get. So, this times 1/det(A), but the point is this thing over here is 0 and why is it 0, where i is not k, why is it 0? Well, that is because we can think of this as the determinant of a matrix in which the i^{th} row is the same as the i^{th} row of A and the k^{th} row is also the same as the i^{th} row, so two rows are same in this.

So, you can think of this as a determinant of a matrix with two rows the same, namely all rows of A are exactly the rows of this matrix except the k^{th} one and the k^{th} one what do we have? We have the i^{th} row. And in the i^{th} row what do we have? We have the i^{th} row. So, two rows are the same and we have seen in the last video that if two rows are the same then the determinant is 0, so that is why this thing is 0.

And from here it will follow that the ik-th entry is also 0, so we will get that $A\frac{1}{det(A)}adj(A)$ is the identity matrix. Similarly, you can flip it around and work with columns and you will get that is also the identity matrix, so that is how we get this formula for A^{-1}.

(Refer Slide Time: 22:54)

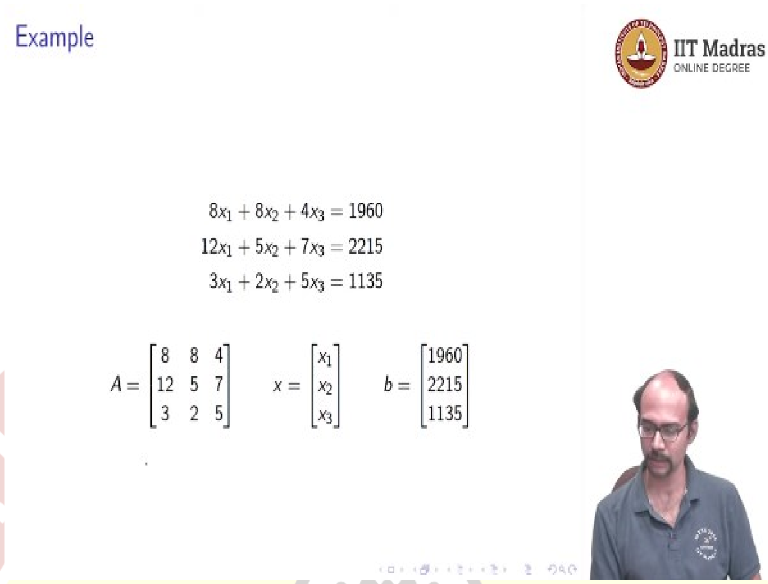


So, now let us get to what we really wanted to do, namely we want to find the solution of a linear, system of linear equations where the coefficient matrix is invertible. So, I will repeat what that means, the coefficient matrix is invertible means that it has an inverse which means the same as its determinant is non-zero that is what we have seen, that if the determinant is non-zero, then it has an inverse and of course, if it has an inverse the determinant is non-zero.

So, let us look at the system of linear equations Ax = b where the coefficient matrix A is an invertible matrix, so let us multiply both sides by A^{-1}, what do we get?

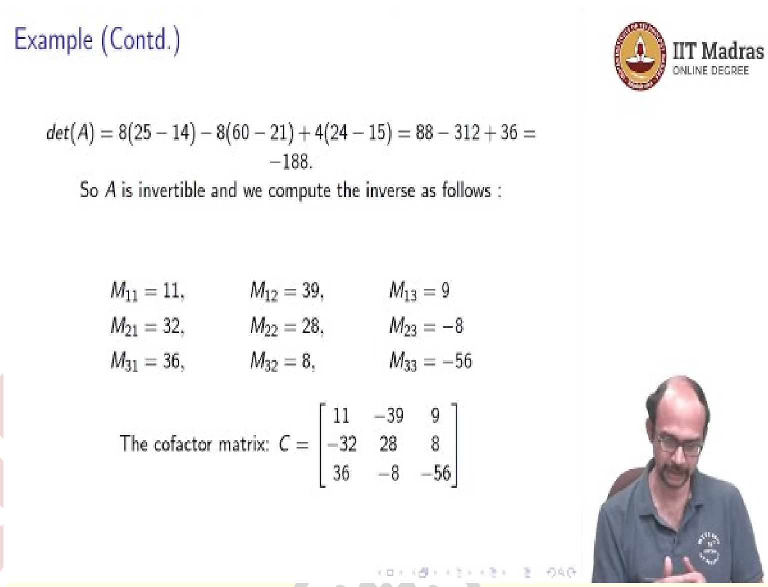
So, $A^{-1}Ax= A^{-1}b$. So, I get the identity matrix, so the $Ix = A^{-1}b$, but identity times anything is that thing itself, so identity times x is x. So, we get $x = A^{-1}b$, so this is the solution. So, we wanted to find the solution and here is the explicit form for the solution.

(Refer Slide Time: 24:31)



So, let us do an example, so here is an example that we actually computed from much, much before, so this was actually something to do with price of rice and dal and so on. So, $8x_1 + 8x_2 + 4x_3 = 1960$, $12x_1 + 5x_2 + 7x_3 = 2215$, $3x_1 + 2x_2 + 5x_3 = 1135$. So, let us form the corresponding system of linear equations, that is this is, this $A= \begin{bmatrix} 8 & 8 & 4 \\ 12 & 5 & 7 \\ 3 & 2 & 5 \end{bmatrix}$, x is the column vector of unknowns and b is the column vector of the constants 1960, 2215 and 1135. So, now the first thing we want to check is what is determinant of A.

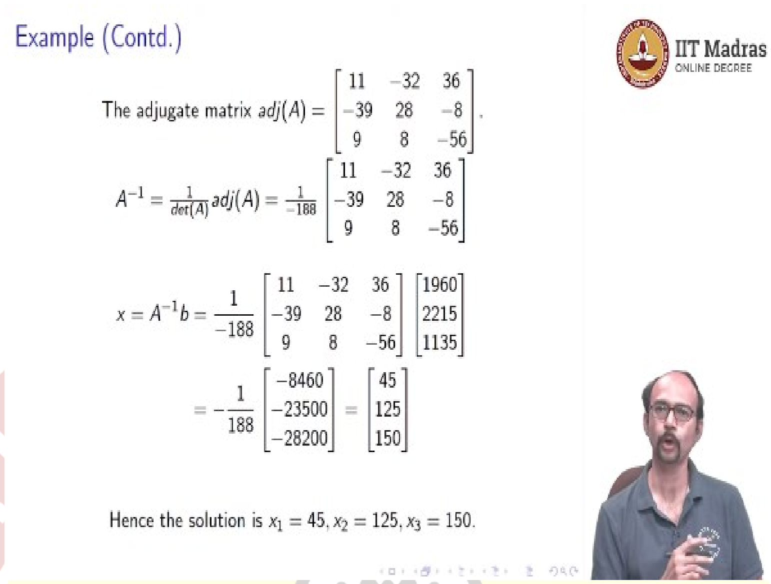
(Refer Slide Time: 25:35)



So, determinant of A we can compute as shown above it gives you -188 which is non-zero, which is really all we care about. So, once it is non-zero that means A is invertible and we know explicitly what the solution is. It is A^{-1} b, so now what we have to do is we have to compute A^{-1} and then we have to multiply it with b. So, how do we compute A^{-1} well?

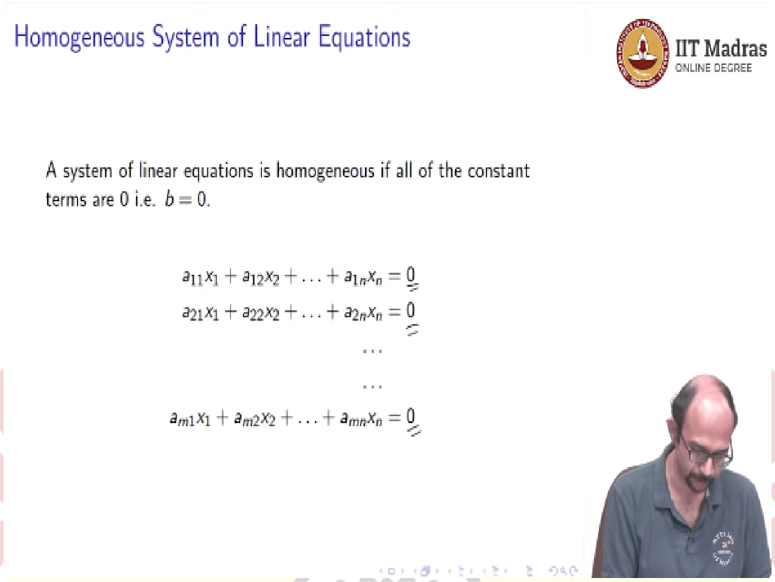
We have to find the minors, this is what the minors are so 11, 39, 9, 32, 28, -8, 36, 8, -56 you will have to do this, I trust now we can all compute determinants and so you will be able to do this. So, just to reiterate to compute the (3, 2) minor you have to hide the third row and the second column and compute the determinant of the remaining matrix. So, from the minors we can get the co-factor matrix, how do we do that? We multiply each of these by $(-1)^{i+j}$, so that is how we get the cofactor matrix.

(Refer Slide Time: 26:40)



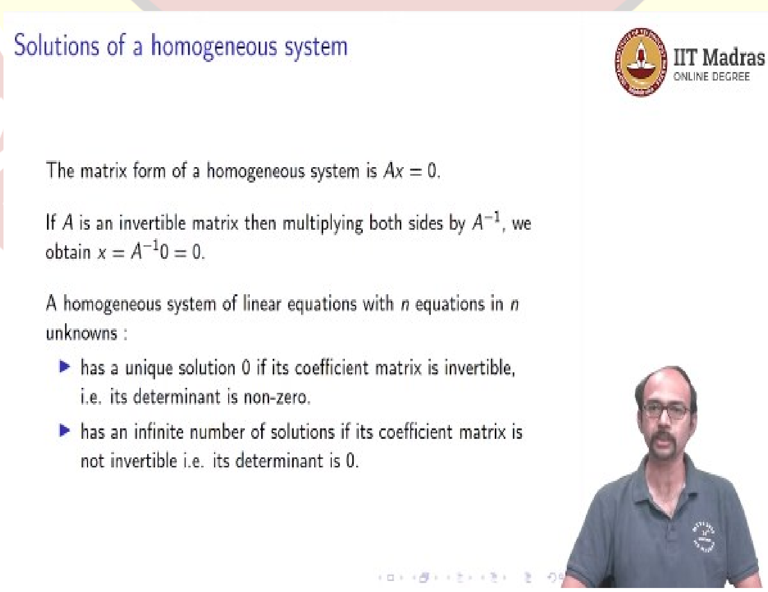
And then we can form the adjugate matrix, the adjugate is the transpose of the cofactor matrix, so that is what we get and then $A^{-1} = \frac{1}{det(A)} adj(A)$. And then what is x? Which is the solution, the solution is A^{-1} b, so you multiply all these and then finally what you get is 45, 125, 150 which means x_1 is 45, x_2 is 125, x_3 is 150. Now, I will ask you to compare this with the solution you got, we worked this out when we did systems of linear equations. So, look at that video and observe that we got the same thing.

(Refer Slide Time: 27:34)



So, finally we want to look at something called the homogeneous system of linear equations. So, what is the homogeneous system of linear equations? When you write it as Ax = b, so now A need not be a square matrix, so you could have m equations in n unknowns but the b is 0, yeah, the constants appearing here, so these are all 0. So, this is 0, all these are 0. So, if these are 0 it is called a homogeneous system of linear equations.

(Refer Slide Time: 28:05)



Now, if it so happens that you have a homogeneous system of linear equations, so that means the matrix form is $Ax = 0$ and this is a square matrix, so A square, so I should have added that over here, so if A is an invertible matrix, then we know exactly what the solution is, it is A^{-1} b but b is 0 here, so that means $x = 0$. So, a homogeneous system of linear equations with n equations in n unknowns, so it has a unique solution 0, if its coefficient matrix is invertible so that is the same as saying its determinant is non-zero and here is something extra, it has an infinite number of solutions if its coefficient matrix is not invertible, that means its determinant is 0.

So, now we know explicitly that if you have n equations in n unknowns, then either and your system is homogeneous, so you are solving $Ax = 0$, if the determinant of A is non-zero then the solution is A^{-1} 0 which is 0. And if the determinant is 0, then you have an infinite number of solutions, so the first part is something we did, the second part is something I am stating and we will do this in the next video. Thank you.