Essential Mathematics for Machine Learning

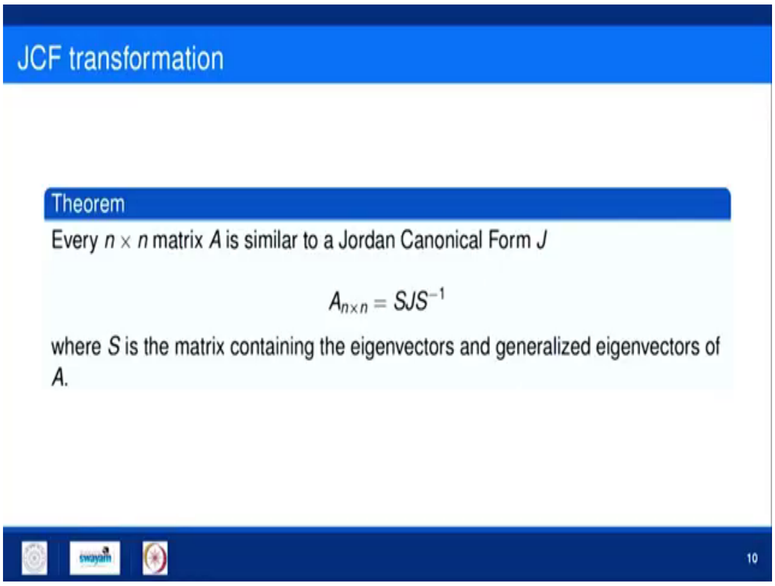
Prof. Sanjeev Kumar
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture - 20

Minimal Polynomial and Jordan Canonical Form

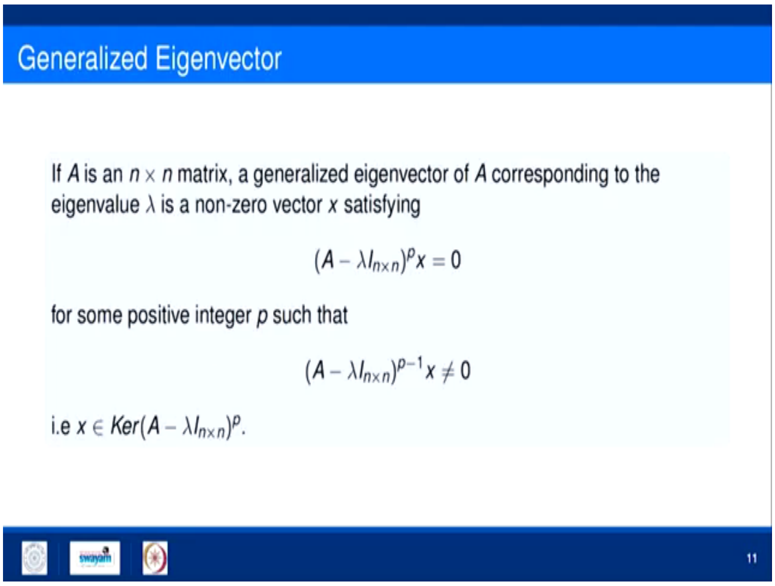
Hello friends. So, welcome to the another lecture on Minimal Polynomial and Jordan Canonical Form of the course Essential Mathematics for Machine Learning. In the last lecture we have learn about minimal polynomial and Jordan canonical form of a matrix, in this lecture we will continue from the previous lecture and we will learn about the Jordan canonical transformation as well as what is the relation of minimal polynomial with Jordan canonical form of a matrix.

(Refer Slide Time: 01:03)



So, I will start this lecture with Jordan canonical form transformation result. That every n by n matrix a is similar to a Jordan canonical form J that is A equals to SJS inverse where S is the matrix containing the eigenvectors and generalized eigenvectors of A and J is the Jordan canonical form of the matrix A. In the last lecture we have learn how to write J from the matrix A in this lecture we will learn how to calculate this matrix S and then what is the relevance of this particular transformation in machine learning.

(Refer Slide Time: 01:53)

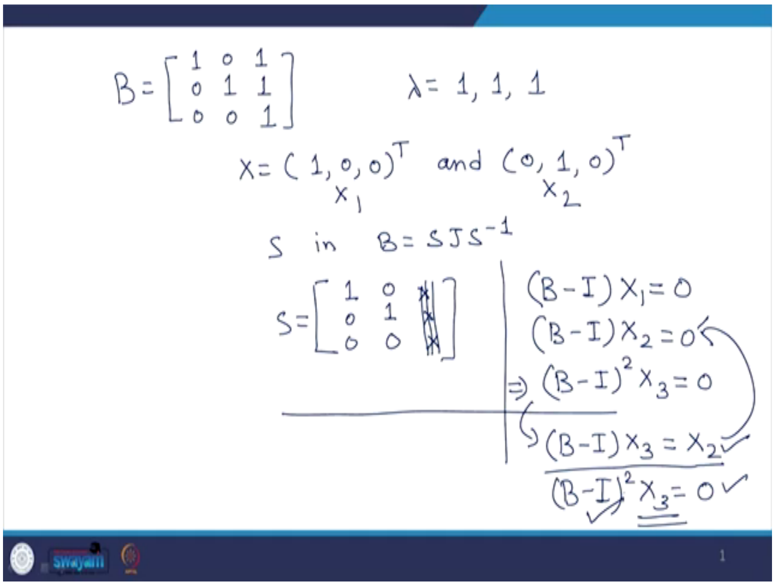


So, for writing S as I told you S is a matrix having columns as eigenvectors of A as well as generalized eigenvectors of A. We know how to calculate eigenvectors, but we do not know how to calculate generalized eigenvector. So, let me define generalized eigenvectors of a matrix A. If A is an n by n matrix a generalized eigenvector of a corresponding to eigenvalue lambda is a non zero vector X satisfying A minus lambda I raised to power p times X equals to

0. For some positive integer p such that A minus lambda I raised to power p minus 1 times X not equals to 0.

It means X is a vector in the null space of a minus lambda I raised to power p what is eigenvector? Eigenvector is a vector in the null space of A minus lambda I whereas, a generalized eigenvector is a vector which is in the null space of A minus lambda I raised to power p; such that it is not a vector in the null space of A minus lambda I raised to power p minus 1 and so, on ok. So, this is the definition of generalized eigenvector.

(Refer Slide Time: 03:38)

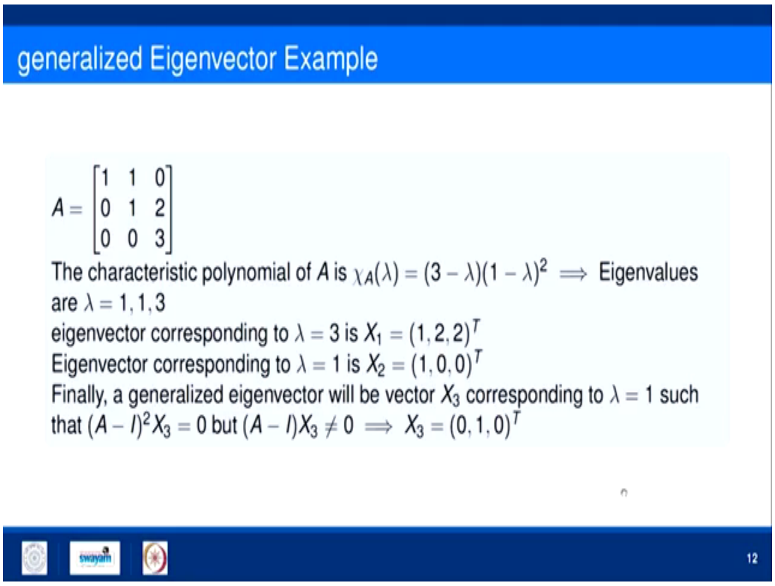


Now, why we need generalized eigenvector? As you know that we have seen a matrix B in the last lecture which is having the form like this 1 0 1, 0 1 1, 0 0 1 and then we have seen that lambda equals to 1 1 1 is the eigenvalue of B. Now eigenvectors comes out to be 1 0 0 transpose and 0 1 0 transpose. So, here geometric multiplicity of lambda equals to 1 is 2 while the algebraic multiplicity is 3 ok. So, when we write S in B equals to SJS inverse then two of the columns of S comes from the eigenvectors 1 0 0 and 0 1 0.

However, what about this column? Because S would be a 3 by 3 matrix because B is 3 by 3. So, these columns will come from the generalized eigenvector, it will be a generalized eigenvector of B now since B minus lambda, lambda is 1. So, B minus I X is having only 2 l I solutions that is X 1 let us say this is my X 1 this is X 2 and B minus I X 2 equals to 0.

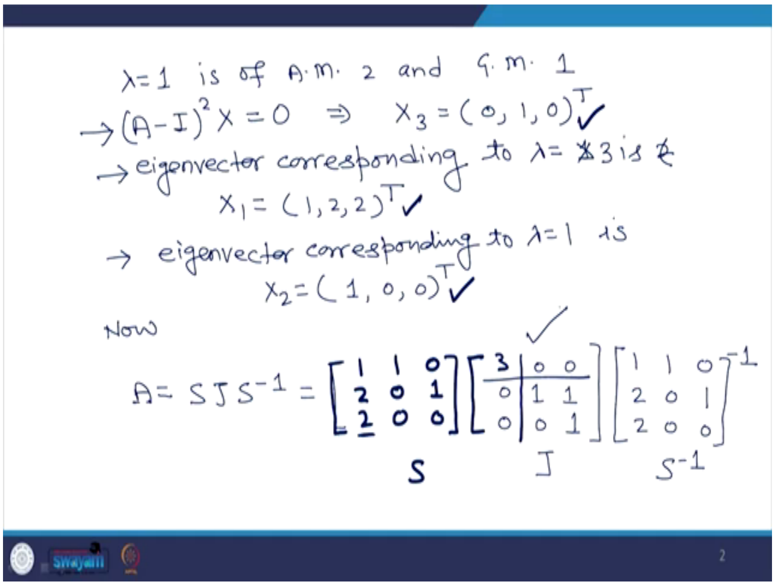
So, third vector I will take a generalized eigenvector which is B minus I square X 3 equals to 0 and how to calculate it? This can be written as B minus I X 3 equals to X 2 by solving this non homogeneous system where X 2 is already known to you or X 1. We multiply both side by B minus I then it will become square here it will become B minus I times X 2 and that is equals to 0. So, hence X 3 is a generalized eigenvector which you can obtained by solving this or by solving this homogeneous system.

(Refer Slide Time: 06:15)



So, now see an example here we are having this matrix it is a 3 by 3 matrix and the eigenvalues are 1 1 and 3. Eigenvalue lambda equals to 1 is having algebraic multiplicity 2 while the eigenvalue lambda equals to 3 will be having algebraic multiplicity 1. Now eigenvector corresponding to lambda equals to 3 comes out to be 1 2 2. Eigenvector corresponding to lambda equals to one comes out to be 1 0 0.

(Refer Slide Time: 06:54)



So, here you can see that lambda equals to 1 is of algebraic multiplicity 2 and geometric multiplicity 1. So, what I will do? I will find out a generalized eigenvector corresponding to lambda equals to 1 that is I will solve this system A minus I whole square X equals to 0. So, by solving this I will get a vector which is 0 1 0.

So, what I got here? Eigenvector sorry lambda equals to 3 is X 1 equals to 1 2 2, eigenvector 1 0 0. Now how to write Jordan canonical transformation of A? So, A equals to S J S inverse what is S here? I will take X 1 that is 1 2 2 then I will be having X 2 1 0 0, X 1 X 2 and now third column comes from the generalized eigenvector that is 0 1 0.

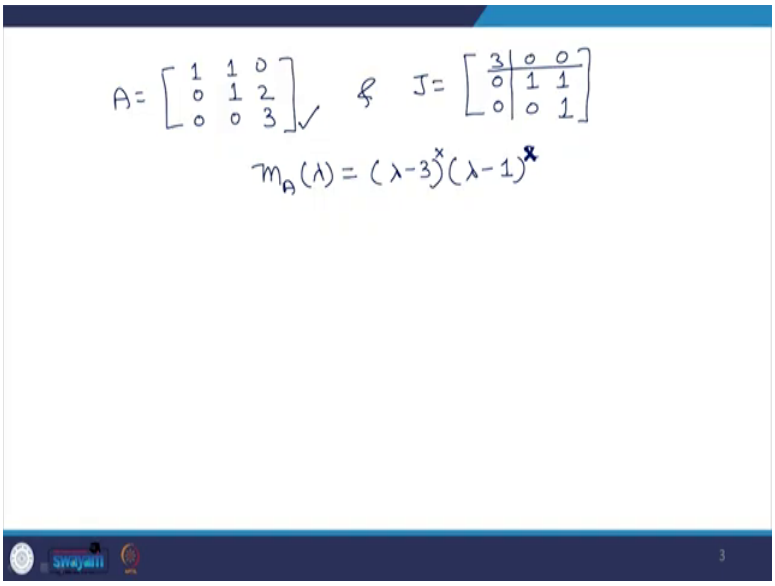
So, this is your matrix S now how to write J? So, you have written eigenvector corresponding to lambda equals to 3 in the first column. So, you write Jordan block corresponding to lambda equals to 3 here. So, one algebraic multiplicity, one geometric multiplicity. So, it will be one Jordan block of size one that is this one.

Now corresponding to lambda equals to 1, algebraic multiplicity is 2 while the geometric multiplicity is 1. So, a Jordan block of size 2. So, it means 3 1 1 0 1 and then 0 0 0 0 and then you will be having this is your J S inverse. So, inverse of 1 1 0 2 0 1 2 0 0. So, this is Jordan canonical transformation of A. Now here you can see if you compare the matrix A here J is a sparse matrix when compared to A and sparsity is very important in machine learning.

So, one of the relevance of this Jordan canonical form in machine learning is instead of you can take J as the sparse version of A and then what you can do? You can perform various type of dictionary learning algorithm or let us say compressive type of thing on J you can go to a smaller version compressed version of J, you can perform some of whatever processing you want there basically in image processing based machine learning and then you can come back you can reconstruct it based on the l 0 optimization that is the sparse norm of optimization and then you can reconstruct back your matrix A with the help of S that is one of the relevance.

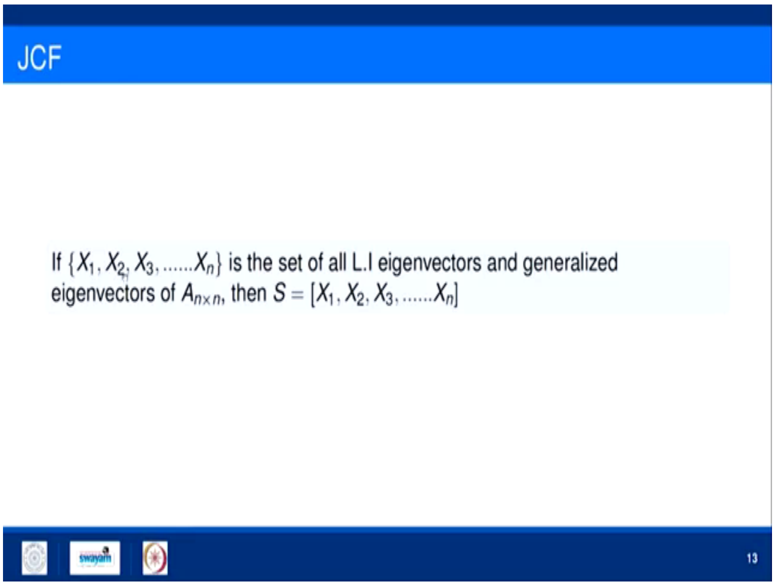
Now this is the case here now I want to show you another important property of this matrix A in terms of Jordan canonical form that is relevance with minimal polynomial. So, you see here if I ask you, what will be the minimal polynomial of A? So, if you see here A is 1 1 0 0 1 2 0 0 3 and J comes out to be 3 0 0 0 1 1 and 0 0 1.

(Refer Slide Time: 12:03)



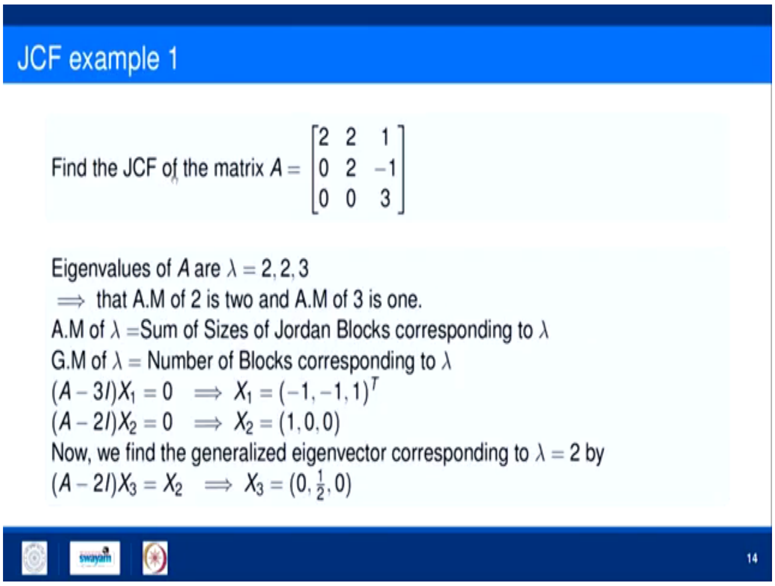
Now if I ask you tell me about minimal polynomial of A. So, minimal polynomial of A I can write from here and it will be lambda minus 3 which is coming from the Jordan block and then from the Jordan block I will write lambda minus 1 square how? So, if this power this power in the minimal polynomial gives you the size of largest Jordan block corresponding to that particular eigenvalue, we will see an example of it.

(Refer Slide Time: 13:13)



So, if X 1 X 2 X n is set of L I eigenvectors and generalized eigenvector, then S can be written in this way that is each eigenvector can be written in the column in the Jordan canonical form which we have seen from the example.

(Refer Slide Time: 13:31)

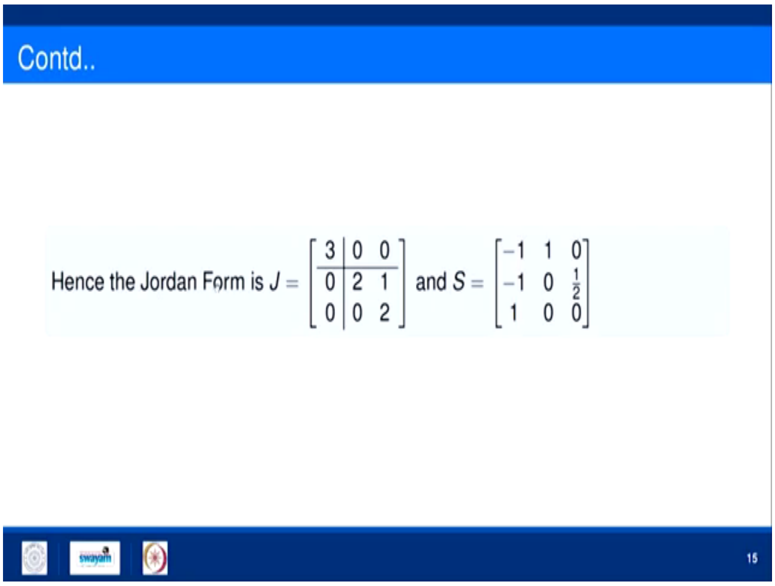


Now, find the Jordan canonical transformation of this matrix A another example. So, matrix is 2 2 1 0 2 minus 1 0 0 3. So, from here we can see eigenvalues are 2 2 and 3 because it is an upper triangular matrix so algebraic multiplicity of lambda equals to 2 is 2 while the algebraic multiplicity of lambda equals to 3 is 1.

Now what algebraic multiplicity tells you? Sum of sizes of Jordan blocks corresponding to lambda. So, what should be the sum of sizes of Jordan blocks corresponding to lambda equals to 2? It should be 2 and geometric multiplicity tells the number of blocks. So, if we calculate here, we see that geometric multiplicity of lambda equals to 2 is 1 and geometric multiplicity of lambda equals to 3 is also 1.

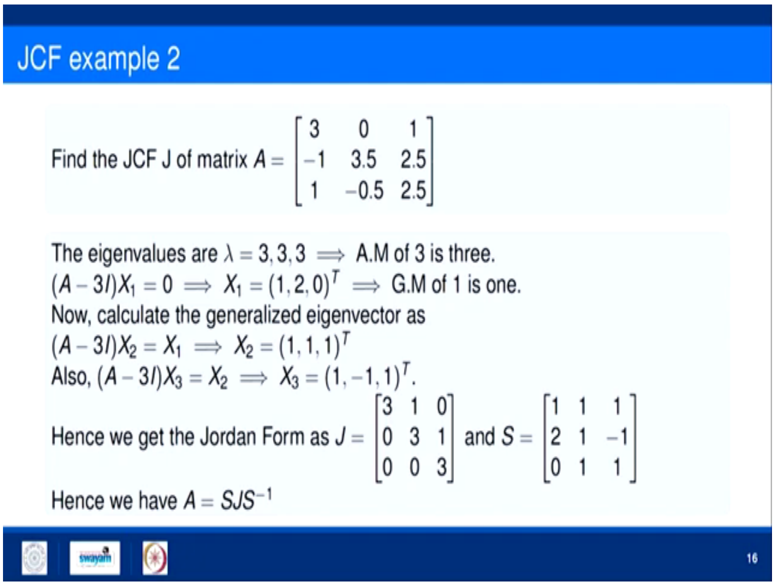
So, you will be having one Jordan block of size 1 corresponding to lambda equals to 3 while one Jordan block because G M is 1 for lambda equals to 2. So, one Jordan block corresponding to lambda equals to 2 of size equals to algebraic multiplicity of lambda equals to 2 that is 2 and then you can find out the third column of this matrix S I calculating the generalized eigenvector corresponding to lambda equals to 2.

(Refer Slide Time: 15:01)



And then by doing this we got the Jordan canonical form.

(Refer Slide Time: 15:05)

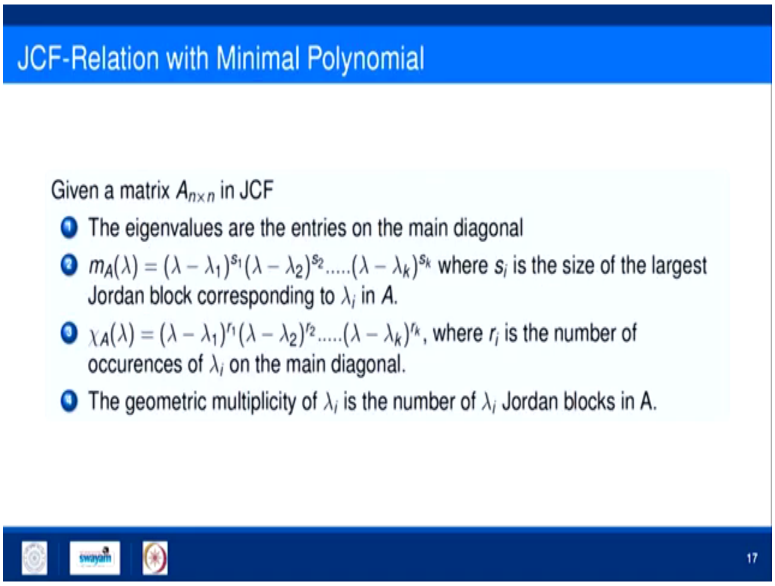


Another example it is very interesting example. So, you see this matrix say again 3 by 3 matrix, I am taking eigenvalue of this matrix comes out to be 3 3 3. So, one eigenvalue is 3 with algebraic multiplicity 3. So, in Jordan canonical form of A there will be Jordan blocks of size total sum will be 3 corresponding to eigenvalue lambda equals to 3. If I calculate the eigenvector corresponding to lambda equals to 3, then A minus 3 I X 1 equals to 0 gives me X 1 equals to 1 2 0. So, hence geometric multiplicity of 1 is 1. Now calculate generalized eigenvector. So, I need to calculate two generalized eigenvectors here.

So, one I am calculating using A minus 3 I square X 2 equals to 0 and that gives me X 2 equals to 1 1 1 and then another one I am calculating A minus 3 I raised to power 3 into X 3 equals to 0 which gives me X 3 equals to 1 minus 1 1. So, here I will be having J only one Jordan block of size 3. So, J becomes 3 1 0 0 3 1 0 0 3 and the corresponding matrix S will be

1 2 0 is the first column 1 1 1 is second column and 1 minus 1 1 is the third column and here A will be S J S inverse.

(Refer Slide Time: 16:50)

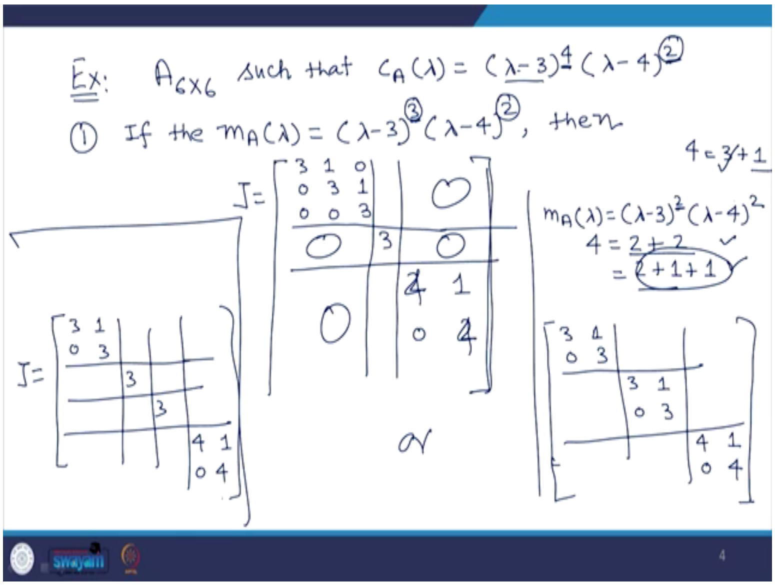


So, what we have seen? If I talk the relation between JCF and minimal polynomial the eigenvalues are the entries on the main diagonal in the Jordan canonical form of a matrix A if the minimal polynomial of A is lambda minus lambda 1 raised to power s 1 lambda minus lambda 2 raised to power s 2 lambda minus lambda k raised to power s k where S i is the size of the largest Jordan block corresponding to lambda i in A.

So, what this power is giving me for each eigenvalue factor? It is giving me the size of larger Jordan block corresponding to that particular eigenvalue. Characteristic polynomial of A is given by this one where r i is the number of occurrence of lambda i on the main diagonal means, r i are the algebraic multiplicity of each lambda I corresponding lambda i the geometric multiplicity of lambda i is the number of Jordan blocks in A.

So, s i is giving me the size of larger Jordan block while the geometric multiplicity is giving me the total number of blocks and if I am having these information's that about s i r i's and geometric multiplicities then I can write the Jordan canonical form of a matrix.

(Refer Slide Time: 18:28)



Let us see an example based on these information. A is a 6 by 6 matrix such that the characteristic polynomial of A is lambda minus 3 raise to power 4 into lambda minus 4 raised to power 2. So, what information we are getting from this that, we are having eigenvalue 3 with algebraic multiplicity 4 and eigenvalue 4 algebraic multiplicity of lambda equals to 4 is 2 now 1.

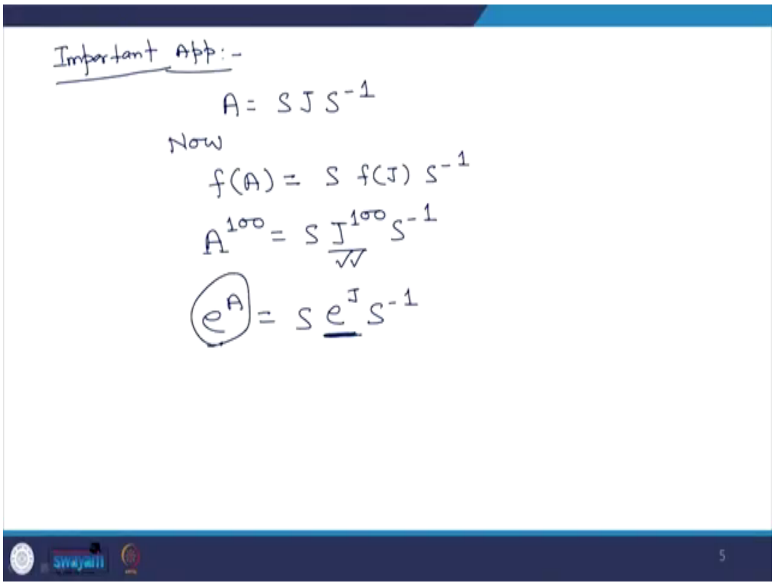
If the minimal polynomial of A is lambda minus 3 raised to power 3 and lambda minus 4 raised to power 2, then what will be the Jordan canonical form? So, let us try to write it what this thing is telling me? The size of larger Jordan block. So, I have the sum of total sizes of the Jordan blocks corresponding to lambda equals to 3 is 4 out of which largest Jordan block is of size 3. So, for largest size is 3. So, only possibility left that another Jordan block of size 1.

So, I can write here 3 1 0 0 3 1 0 0 3 that is one of the Jordan block that is corresponding to this factor, then another of size 1. So, it means what is geometric multiplicity of lambda equals to 3? 2 that is the total number of Jordan blocks corresponding to lambda equals to 3, then you are having lambda minus 4 here total size is 2 largest block size is 2. So, the only possibility there is only one block of size 2.

So, this is the Jordan canonical form of A; however, if I am having this minimal polynomial is. So, let me do it here if m A lambda is giving you lambda minus 3 square lambda minus 4 square, then now total size corresponding to lambda equals to 3 is 4 out of which largest size is 2. So, what I am having? Total size for largest is 2. So, what possibility left? Another maybe of size 2 or it is 2 plus 1 plus 1.

So, here we are having two possibilities while there will be a Jordan block corresponding to lambda equals to 4 of size 2. So, what will be the Jordan canonical form in this case? If I consider this particular thing then 3 1 0 3 that is first Jordan block 3 1 0 3 and then 4 one sorry it is 4 here 4 1 0 4 or if I take this particular case then in this case J becomes 3 1 0 3 3 3 and then 4 1 0 4 ok. So, in that way you can write the Jordan canonical form of a given matrix based on the information given to you about characteristic polynomial and minimal polynomial.

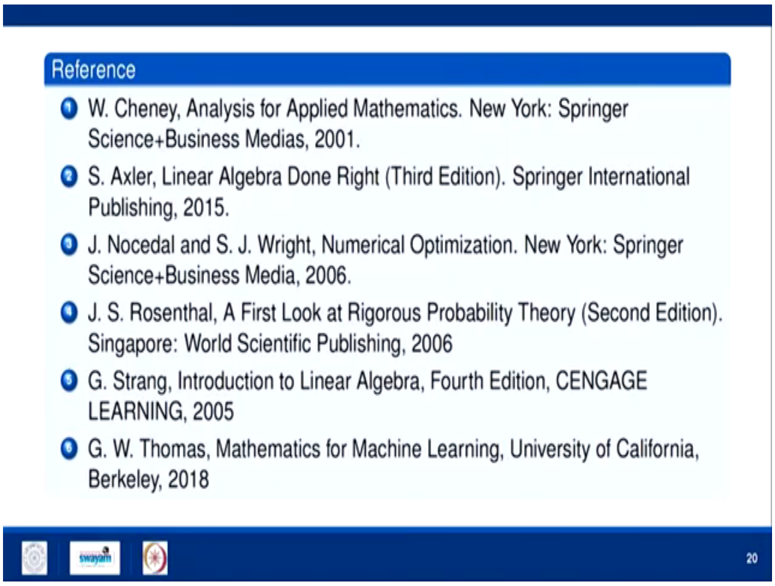
(Refer Slide Time: 23:21)



Another important aspect I will say the applications, here is of this similarity transformation is calculating the functions of the linear transformations or functions of the matrix. So, you know that I can write a matrix A equals to S J S inverse. Now a certain class of functions f if you are calculating this function of A, it becomes S times f J into S inverse similar to diagonalization. So, for example, I need to calculate A raised to power 100 then which become S J raised to power 100 into S inverse calculating J raised to power 100 of a sparse matrix is quite easy when compared to calculating the 100 power of A dense matrix.

Another if you need to calculate e raised to power A then S e raised to power J S inverse and now when you will calculate this e raised to power J you have to calculate it Jordan block wise and that will be quite easy when compared to e raised to power A with matrix A. Similarly we can have trigonometric functions sin a cos a those are having some of infinite terms and you can calculate here it very easily. So, this is another application of this important this particular similarity transformation.

(Refer Slide Time: 25:15)



These are the references for the last two lectures means this lecture and previous lecture, I hope you have enjoyed this lecture.

Thank you very much.