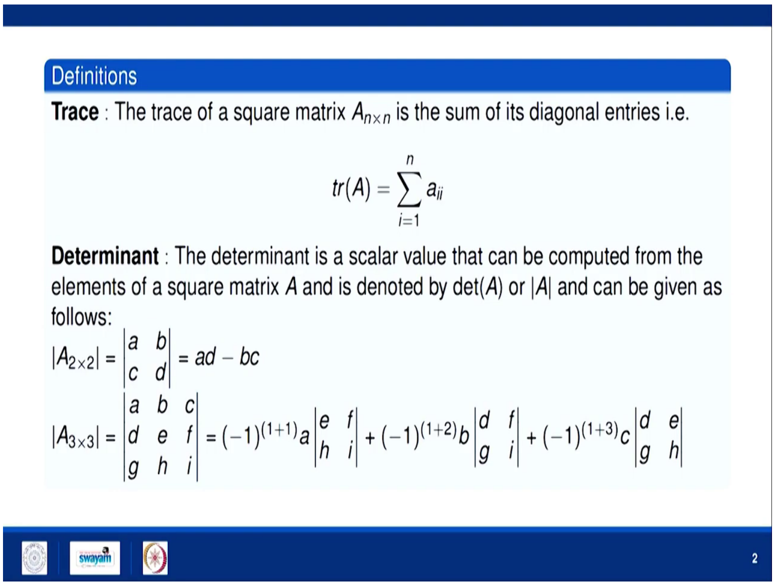
Essential Mathematics for Machine Learning
Prof. Sanjeev Kumar
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture - 08
Eigenpairs and Properties

Hello friends. So, welcome to the lecture number-8 of this course on Essential Mathematics for Machine Learning. So, in last lecture, we have discussed about orthogonal complements and projection transformation. Today, we are going to discuss about Eigenpairs means eigenvalues and eigenvectors associated with a square matrix and their Properties.

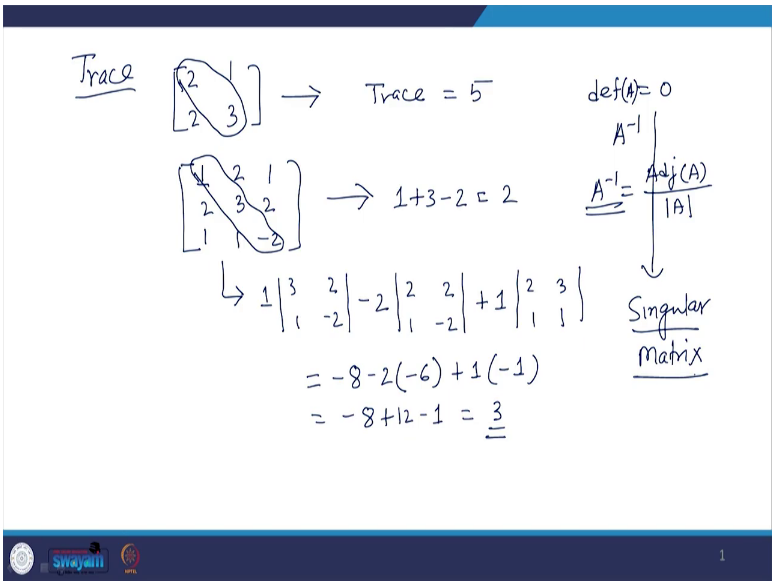
So, eigenpairs in terms of eigenvalues and eigenvectors are very crucial in machine learning, because you know we have to transform feature space, we have to make several analysis; and for doing those kind of thing, we need the concept of eigenvalues and eigenvectors. Even though in dimensional reduction algorithm like principal component analysis, linear discriminant analysis, although all those kind of algorithms are based on eigenvalues and eigenvectors only.

(Refer Slide Time: 01:29)



So, we will start this lecture with few definitions. So, first I am going to define trace of a matrix. So, the trace of a square matrix A is the sum of its diagonal entries that is trace of A, if A is n by n matrix then it is sum of all diagonal elements.

(Refer Slide Time: 01:53)



So, for example, if you are having a matrix let us say 2 by 2 matrix 2 3 1 2. So, what is the trace? Sum of diagonal elements. So, these are the diagonal elements. So, trace of this matrix is 5. Similarly, if you are having a 3 by 3 matrix, let us say 1 2 1 2 3 2 1 1 minus 2. So, what is the trace of this matrix, just sum of the diagonal elements. So, it is 1 plus 3 minus 2. So, it is 2. So, in that way we can define the trace of a matrix.

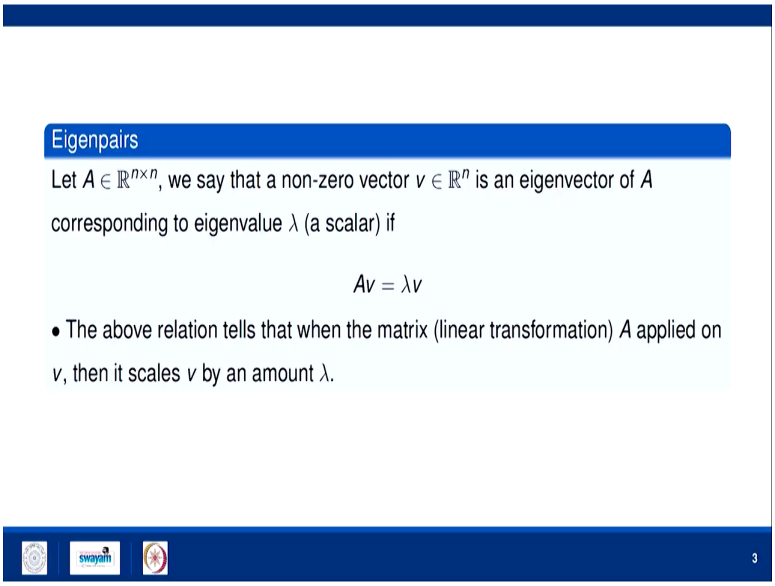
My next definition is determinant. It is very basic definition we used to have it in plus 2 mathematics, but again just I am recalling it. The determinant is a scalar value that can be computed from the elements of a square matrix A. And it is denoted by the det A or by this symbol. If you are having a 2 by 2 matrix let us say elements are a, b, c, d, then the determinant of this matrix is just a d minus b c. If you are having a 3 by 3 matrix, then you can calculate determinant using these cofactors of this matrix.

So, it is you take a, and then just eliminate this row and this column, and then just find the determinant of this 2 by 2 matrix. So, a times determinant of this 2 by 2 matrix, then again you take b, but with minus symbol. So, it alternating plus minus plus minus plus minus like this. So, if elements is a i j where i is 1 to n and j is 1 to n, then for each time we will be having sign like minus 1 raised to power i plus j. So, we when we will take b, so b is 1 2 so or power of minus 1, so I will take minus b and then I will leave this column and this row; and whatever rest is there I will take the determinant of that.

So, if we are having a matrix like this, so this matrix itself. So, if I calculate the determinant of this, so I will take this 1 here, and then determinant of 3 2 1 minus 2, then I will take 2 with minus symbol sign, and then I will leave this row and this column. So, determinant of 2 2 1 minus 2 then I will take plus 1, and then I will leave third column and first row and it is 2 3 1 1.

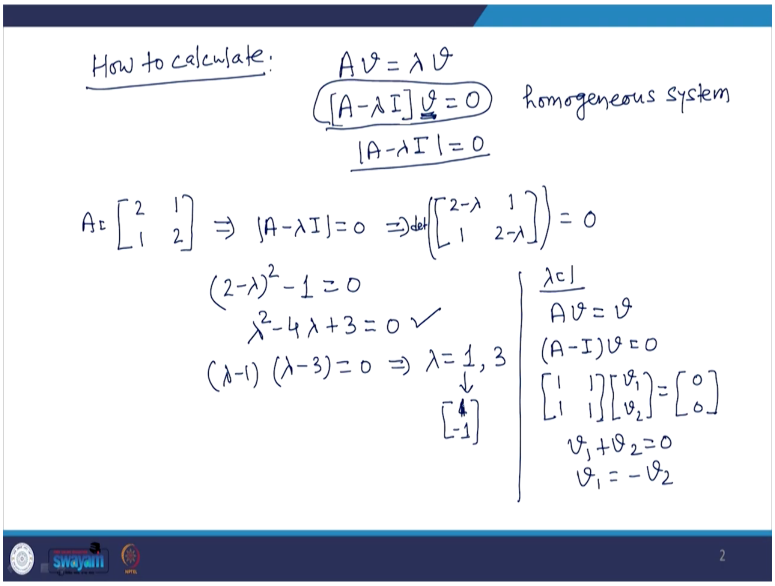
So, basically it will become minus 6 minus 2 minus 8, it will become minus 2 times minus 4 minus 2 minus 6 plus 1 and minus 1 determinant of this. So, it is minus 8 plus 12 minus 1, so determinate comes out to be 3. So, similarly we can calculate determinant of any square matrix. If determinant is 0 of a matrix A, then A inverse does not exist, because you know A inverse can be defined by adjoint of A upon determinant of A. So, determinant of A is 0. So, this quantity is now defined. So, if determinant of A is 0, we say the matrix is singular matrix.

(Refer Slide Time: 06:14)



So, now, come to the definition of eigenvalue. So, let A be a real matrix of size n by n, we say that a non-zero vector. So, please see this non-zero vector v which is n-dimensional vector is an eigenvector of a corresponding to the eigenvalue lambda which is a scalar coming from the field because here A is a linear transformation from n dimensional vector space to n-dimensional vector space, and then lambda is coming from the associated field to those vector spaces. If, so v is an eigenvector corresponding to eigenvalue lambda if a v equals to lambda v. So, what this relation is telling you? So, if you apply this linear transformation on the vector v, it will scale the vector v by an amount of lambda.

(Refer Slide Time: 07:16)



So, we can calculate it using this relation. So, how to calculate? So, we are saying that v is a non-zero vector such that A v equals to scalar lambda times v you take this side left hand side everything. So, I will be having a minus lambda I into v equals to 0. So, now, it is a homogeneous system of linear equation. So, this system will be having non-zero solution that is non-zero v only when rank of or determinant of A minus lambda I is 0.

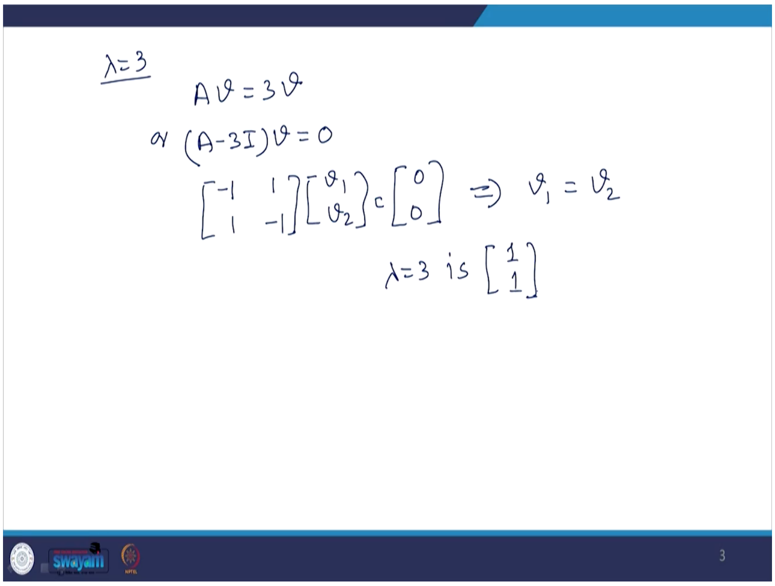
So, if you find this, it will give you if the size of A is n by n, it will give you n-degree polynomial that polynomial is called characteristic polynomial of A. And the roots of that characteristic polynomial will give you the eigenvalues. Once you are having eigenvalues, then by solving this homogeneous system corresponding to each eigenvalue will give you the eigenvector v.

So, for example, if we take a matrix A which is 2 1 1 2, then how to find out its eigenvalues and eigenvector as I told you for finding the eigenvalues I will be having a minus lambda I equals to 0. So, it means A minus lambda I will become 2 minus lambda 1 1 2 minus lambda ok, and determinant of this matrix equals to 0. So, now, find out the determinant of this. So, it will become 2 minus lambda whole square minus 1 equals to 0; or if I simplify it lambda square minus 4 lambda plus 3 equals to 0. So, this is the characteristic lambda square minus 4 lambda plus 3 is the characteristic polynomial of this matrix A.

Now, we will solve this equation. So, the solution of this equation will be lambda minus 3 and lambda minus 1 equals to 0. So, it is giving me 2 solutions lambda equals to 1 and 3. So, the eigenvalue of A are 1 and 3. Now, I will find out the eigenvector corresponding to lambda equals to 1. So, for lambda equals to 1, I will be having a v equals to lambda times v. So, this is A minus I into v equals to 0. So, what is A minus v 1 1 1 1? And let us say this v is simply v 1 and v 2 having these 2 components. So, this equals to 0, 0. So, I got this homogeneous system.

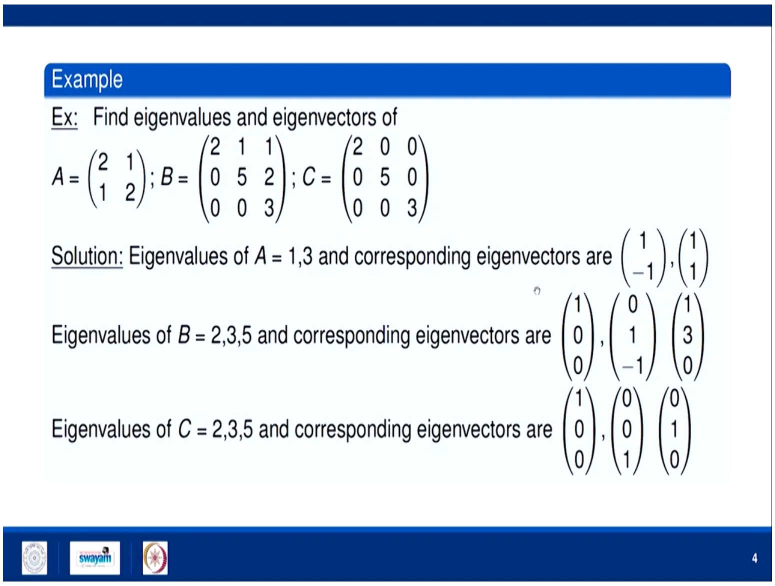
Now, what is the solution? This is saying me that v 1 plus v 2 equals to 0. So, it is saying that v 1 equals to minus v 2. So, if I fix v 1 as 1, I will get v 2 as minus 1. So, eigenvector corresponding to lambda equals to 1 is 1 and minus 1. Similarly, we can find eigenvector corresponding to lambda equals to 3.

(Refer Slide Time: 11:34)



So, for finding the eigenvector corresponding to lambda equals to 3, I will be having A v equals to 3 v, or A minus 3 I, where I is the identity matrix of size 2 by 2 into v equals to 0. So, from here what I will get 2 minus 3 is minus 1 1 1 minus 1. And again v 1 v 2 equals to 0 0. So, from here I will get a relation v 1 equals to v 2. So, if I take v 1 equals to 1, I got v 2 also 1. So, eigenvector corresponding to lambda equals to 3 is 1 and 1. So, in that way, we can find out eigenvalues and eigenvectors of a matrix.

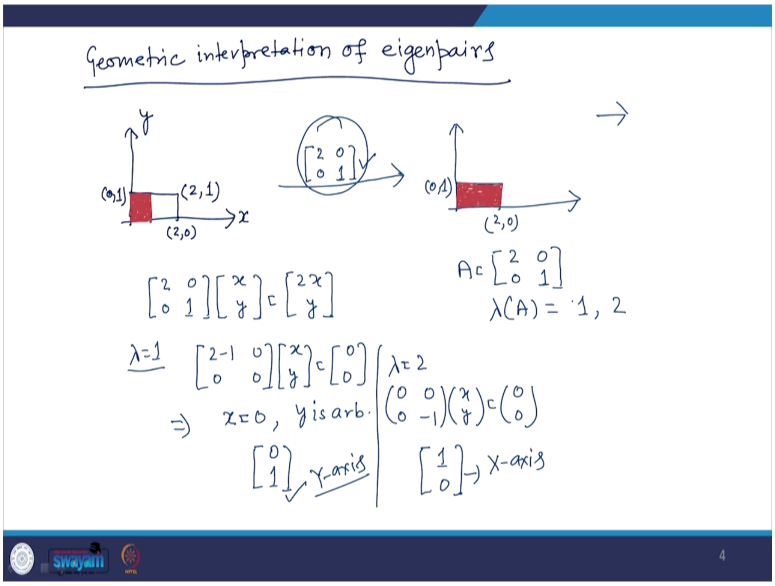
(Refer Slide Time: 12:34)



So, similarly, if you are having this matrix B 2 1 1 0 5 2 0 0 3 actually if you are having a upper triangular matrix or lower triangular matrix, or a diagonal matrix like C, then eigenvalues are just diagonal entries that is coming from the A minus determinant of A minus lambda I equals to 0. So, no need to calculate in these special cases, and you can directly write the eigenvalues, and then you can find out the corresponding eigenvectors.

So, for B eigenvalues are the diagonal entries that is 2 3 and 5, and corresponding eigenvectors are 1 0 0, 0 1 minus 1, 1 3 0. Similarly, you can say about this matrix C. So, eigenvalues are diagonal entries. And if you are having diagonal matrix, then eigenvectors are just standard basis.

(Refer Slide Time: 13:43)



Geometrically, how we will understand this concept of eigenvalues and eigenvector? Geometric interpretation of eigenpairs. So, just see like this. Suppose, I am having the geometry in x y plane. So, we are having this rectangle here. So, let us say it is of length 2. So, I am saying x-axis, and y-axis. So, length is 2; and this is 0, 1. So, this point is 2 and 1. Now, suppose in this, what we are having, this rectangle is half covered by some red object. So, this red object is nothing just a square of length 1 having 1 corner at origin.

Now, what I am doing? I am applying a linear transformation or matrix let us say 2 0 0 1 on this red cover area. So, how this red cover area will transform this object? So, I am applying 2 0 0 1 on x y, so it will become simply 2 x and y. So, what it is saying you that whatever x you are having, it will become double; and whatever y you are having it will remain as such. So, what will be the output? So, after transformation, you will be having same rectangle 2, 0, 0, 1 and then now it will be covered completely by the red object.

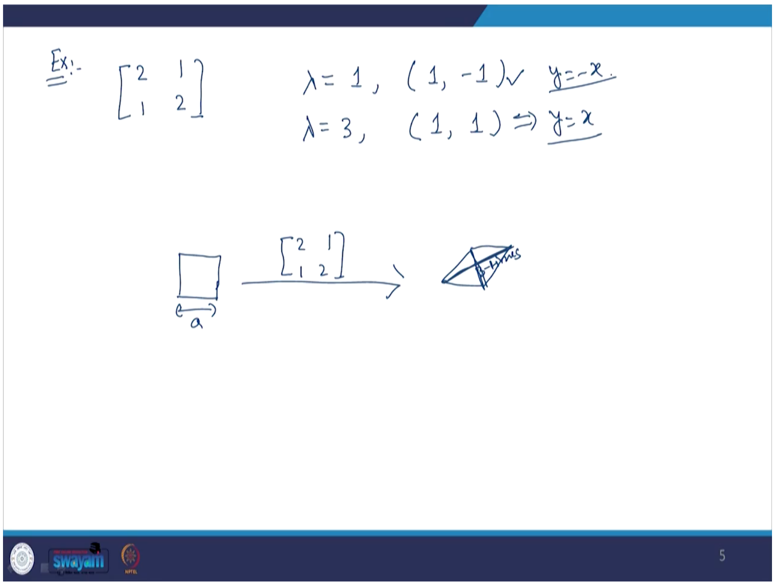
Now, what I want to tell you that now see about the eigenvalues and eigenvectors of this transformation. So, I am having a transformation matrix A which is 2 0 0 1. So, what will happen now eigenvalues of this is 1 and 2. Now, what is eigenvector corresponding to lambda equals to 1? So, for lambda equals to 1, I will be having 2 minus 1 0 0 0 and then x y equals to 0 0. So, what it is giving you x equals to 0, and y is arbitrary. So, y is arbitrary. So, it means I can have an eigenvector x I can fix a 0, and y is 1.

Similarly, for lambda equals to 2, I am getting the system 0 0 0 minus 1, and x y 0 0. So, it is saying y equals to 0. So, I can have x as 1 and y as 0. So, what is this it is nothing just y-axis because x component is 0. Similarly, what is this, x-axis. Now, see I am having eigenvalue of this transformation as 1 and 2; corresponding to 1, I am getting y-axis; corresponding to 2, I am getting x-axis. Now, relate these eigenvalues and eigenvectors to this change which this transformation have been made in this region.

So, from here you will observe what I am getting? In the x direction, I am getting double change, it becomes double; and along y-axis, there is no change. So, from this what I can conclude that the eigenvalues and eigenvectors of this particular transformation are characterized by this change, where eigenvectors are giving the direction in which change has been made and eigenvalues are giving the amount of change. So, for eigenvalue 2, I am having x-axis. So, along x-axis, I am having 2 times whatever earlier.

Now, I am having 2 times of that. Along y-axis eigenvalue is 1, so there is no change. So, in that way, we can characterize or we can geometrically interpret eigenvalues and eigenvector.

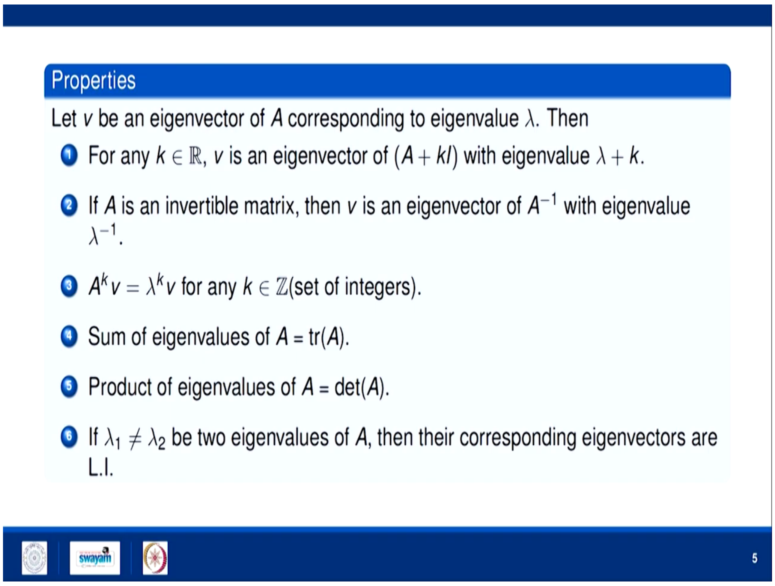
(Refer Slide Time: 19:35)



Just like earlier example which we have shown in first in the beginning of this lecture that is 2 1 1 2. So, we got the eigenvalue as lambda equals to 1 and corresponding eigenvector is 1 and minus 1, and then lambda equals to 3 and corresponding eigenvector is 1 and 1. So, now, again if I am having this kind of rectangle here, how this matrix if I apply this matrix on this rectangle 2 1 1 2, how it will change to this square.

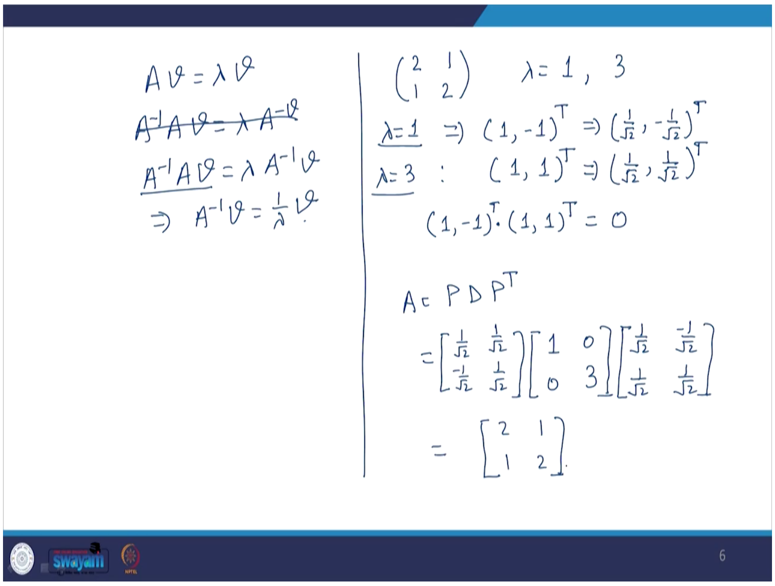
So, it will make it like this. So, along y equals to x, so this is the line y equals to x, I am having 3 times change. So, whatever dimension if it is a, it will becomes 3 a; and there is no change along this direction that is 1 and minus 1 that is y equals to minus x line. So, in that way, we can characterize eigenvalue and eigenvectors for the linear transformations.

(Refer Slide Time: 20:56)



Now, we are going to discuss some of the important properties of eigenvalues and eigenvectors. So, let v be an eigenvector of A corresponding to eigenvalue lambda. Then for any scalar k v v is an eigenvector of A plus k I with eigenvalue lambda plus k. So, what I want to say if you add k times I to A, then eigenvalue will become lambda plus k, and eigenvector will remain same as v. This you can easily verify with the relation A v equals to lambda v. Another important properties if A is an invertible matrix that it is having the non-zero determinant, then v is an eigenvector of A inverse with eigenvalue lambda inverse means 1 by lambda.

(Refer Slide Time: 22:08)

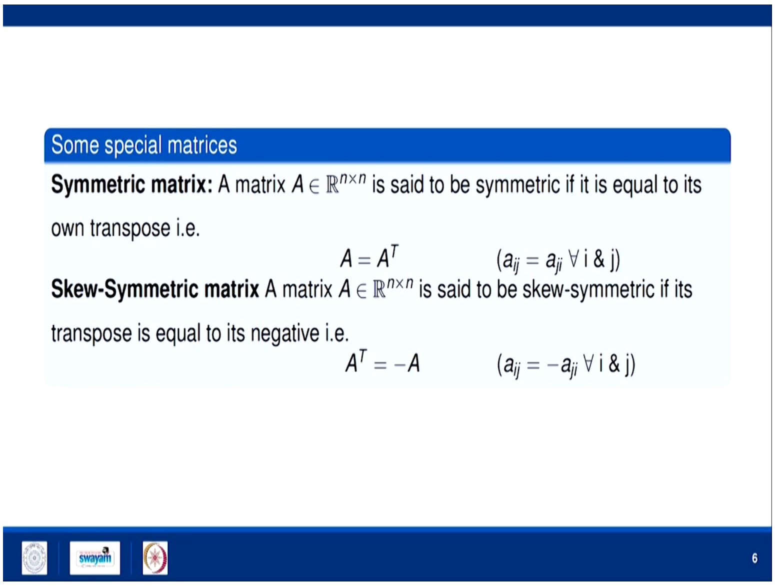


So, if v is an eigenvector, so if A v equals to lambda v, then A inverse v will become 1 upon lambda v. And that you can easily verify from the earlier fact that you are having A v equals to lambda v u multiplied by A inverse both side. So, A inverse A v equals to lambda A A inverse A v equals to lambda A inverse v and this gives you A inverse v equals to 1 upon lambda v, because it will become I v, I v is v only I will take lambda this side and so on. In fact, if A v equals to lambda v then a raise to power k v equals to lambda raise to power k v, for any k belongs to z that is the set of integers either positive or negative.

Sum of eigenvalues of A equals to trace of A. So, what I want to say that sum of the eigenvalues equals to the sum of the diagonal entries of the matrix, and that you can verify from the example which we have taken here eigenvalues are one entity the eigen, so 1 plus 3 is 4 which is equal to 2 plus 2; similarly 2 3 5, 2 2 5 3 2 5 3. So, it can be verified easily.

The product of eigenvalues of A equals to determinant of A. So, what I want to say if a matrix is having 0 eigenvalue, then the determinant is 0, and matrix is not invertible that is it is singular matrix; vice versa we can say if determinant 0 at least 1 of the eigen value of that matrix will be 0. If I am having a matrix A and two eigenvalues are distinct that is lambda 1 not equals to lambda 2, then eigenvectors corresponding to these two matrix, these two eigenvalues will be linearly independent.

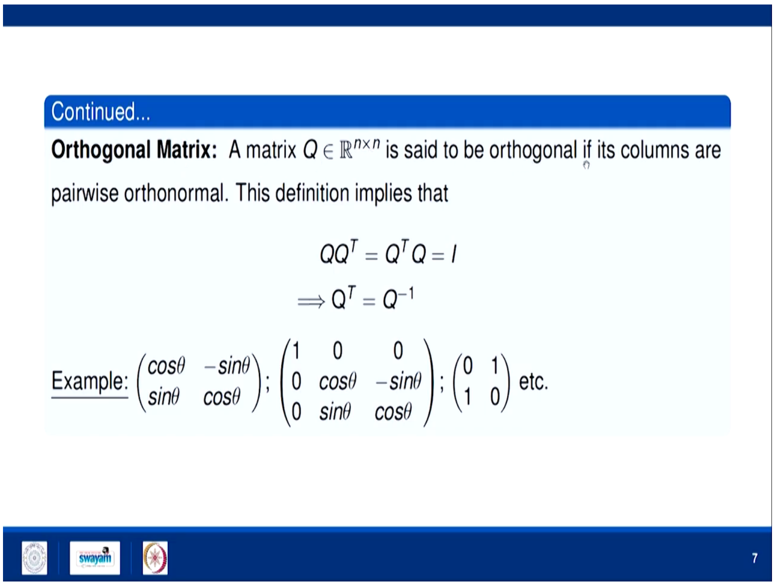
(Refer Slide Time: 24:24)



If I am having a matrix A which is a square matrix of size n by n, and if I am having A equals to A transpose that is if you take the transpose you interchange the row and columns of A, then you are getting the same matrix then such a matrix is called symmetric matrix, like the matrix 2 1 1 2 we have taken in the example.

Similarly, if you are having a skew symmetric means a matrix n by n such that A T equals to minus A, then A is called a skew symmetric. And if you observe from this relation, all the diagonal entries of a skew symmetric matrix will be 0. So, hence space of a skew symmetric matrix will be 0.

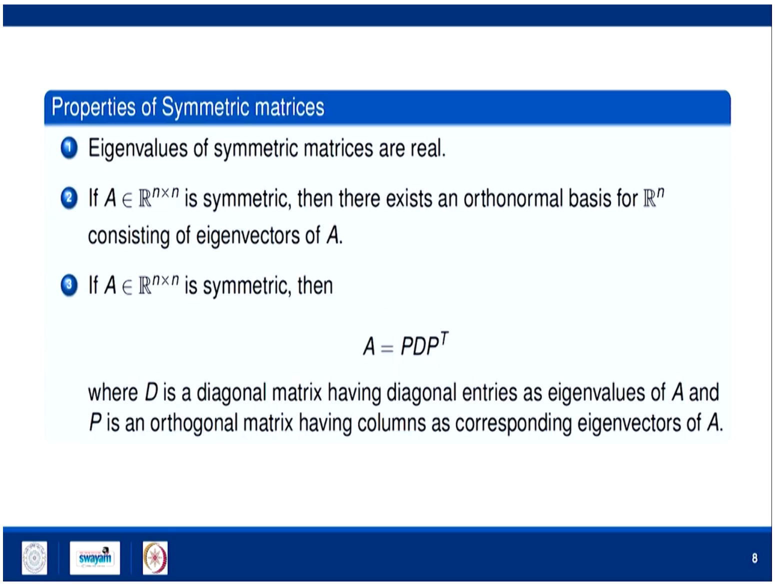
(Refer Slide Time: 25:12)



Another important matrix is orthogonal matrix A matrix Q belongs to R n by n is said to be orthogonal if its columns are pair wise orthonormal. This definition implies Q into Q transpose equals to Q transpose into Q equals to I, or from this I can drive that transpose equals to inverse. So, if the transpose of a matrix equals to inverse means if I am having a orthogonal matrix, then inverse of that matrix will be the transpose itself. Some of the example of orthogonal matrices are rotation matrices.

Similarly, you can say about this matrix this is a matrix rotation matrix in three-dimensional plane, where rotation has been taken about x-axis. Another example is this particular matrix it is again an example of orthogonal matrix.

(Refer Slide Time: 26:15)



The eigenvalues of symmetric matrices are real. Similarly, eigenvalues of a skew symmetric matrix are purely imaginary or 0. If A is n by n symmetric then there exist an orthonormal basis for R n consisting of eigenvectors of A, so that is the spectral decomposition of the matrix A. What I want to say that if a is a symmetric matrix of order n by n, you can always find an orthogonal eigenvectors of A.

And as I told you if you are having those n orthogonal eigenvectors you can make them orthonormal just by dividing each vector by the norm of each one their respective norm, and then I can always write a equals to PDP transpose. Where D is a diagonal matrix having diagonal entries as eigenvalue of A, and P is an orthogonal matrix having columns as corresponding eigenvectors of A.

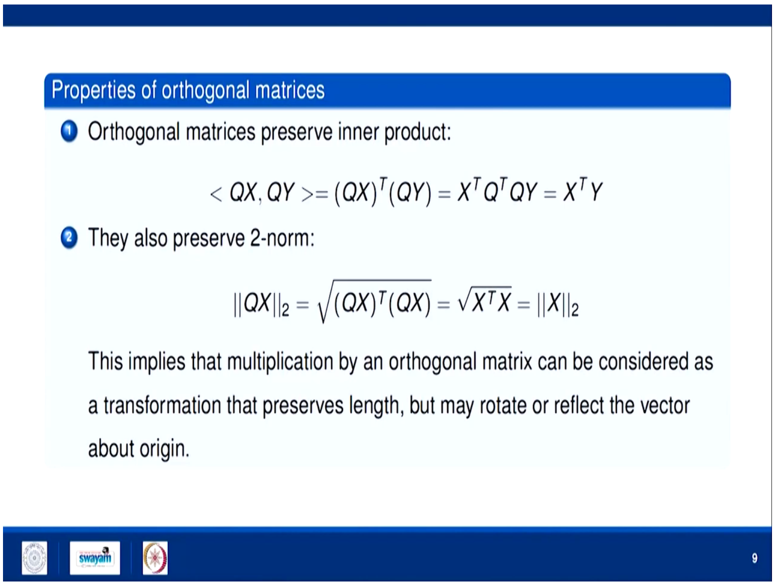
So, for example, we have taken the matrix in the beginning 2 1 1 2. So, it is a symmetric matrix. Now, what was the eigenvalues lambda equals to 1 and 3. What was the eigenvectors corresponding to what is the eigenvector corresponding to lambda equals to 1, that is 1 and minus 1. And what is the eigenvector corresponding to lambda equals to 3 that was 1 and 1. Now, if you see these two eigenvectors are orthogonal, if you take the dot product of 1 minus 1 transpose with 1 1 transpose, this comes out to be 0, 1 minus 1 equals to 0. However, they are not orthonormal. So, make them orthonormal.

So, how? We have to divide it by the length of this vector. So, it will become 1 by root 2 minus 1 by root 2. Similarly, this will become 1 by root 2, and 1 by root 2. Now, what I want to say that A equals to P D P transpose. So, how you will make this matrix P? So, P is nothing just write these orthonormal eigenvectors of A is the columns of P. So, 1 by root 2 minus 1 by root 2 that is the first eigenvector I have written here and then 1 by root 2 and 1 by root 2.

What is D? D is the diagonal matrix having the eigenvalues of A. So, what column I have taken first 1 by root 2 minus 1 by root 2 that is the eigenvector corresponding to lambda equals to 1. So, write 1 here 0 0 3. And then P inverse generally in diagnolization we have P inverse, but here since P is orthogonal, so P inverse equals to P transpose. So, I will be having 1 by root 2 minus 1 by root 2 1 by root 2 and 1 by root 2.

So, if you see now product of these three matrices, it comes out to be 2 1 1 2 that is your matrix A. So, for a symmetric matrix, we can always a this kind of decomposition means we can write A as the product of three matrices where P is an orthogonal matrix and D is a diagonal matrix. In coming lectures, we will see another generalization of this kind of decomposition that is singular value decomposition that is applicable for rectangular matrices also.

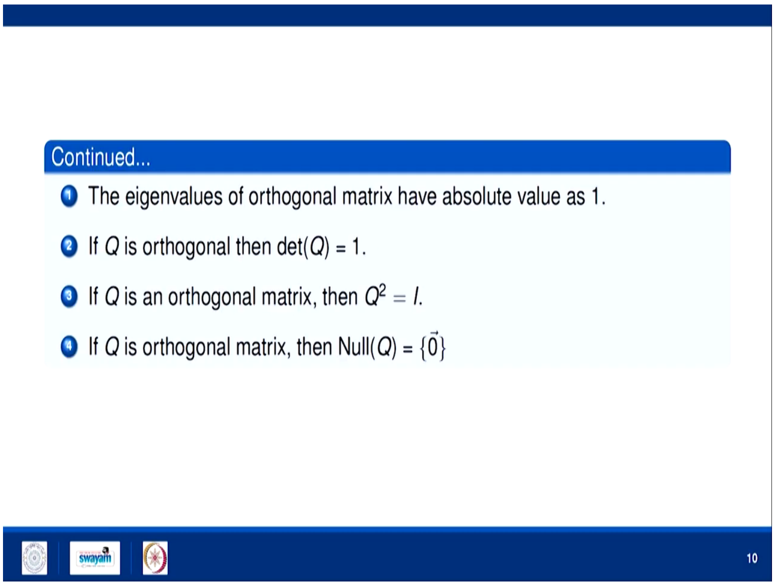
(Refer Slide Time: 31:06)



Another very beautiful property of orthogonal matrices or orthogonal transformations are that they preserve inner product that is if you take if you are having 2 vectors x and y, if you multiply pre multiply Q with X as well as with Y then inner product of Q X Q Y will become Q X transpose into Q I that is X transpose Q T Q Y. Now, Q is an orthogonal matrix. So, Q T Q will become Y. So, it becomes X transpose Y, and which is nothing just the inner product of X and Y. So, inner product of Q X Q Y equals to inner product of X Y.

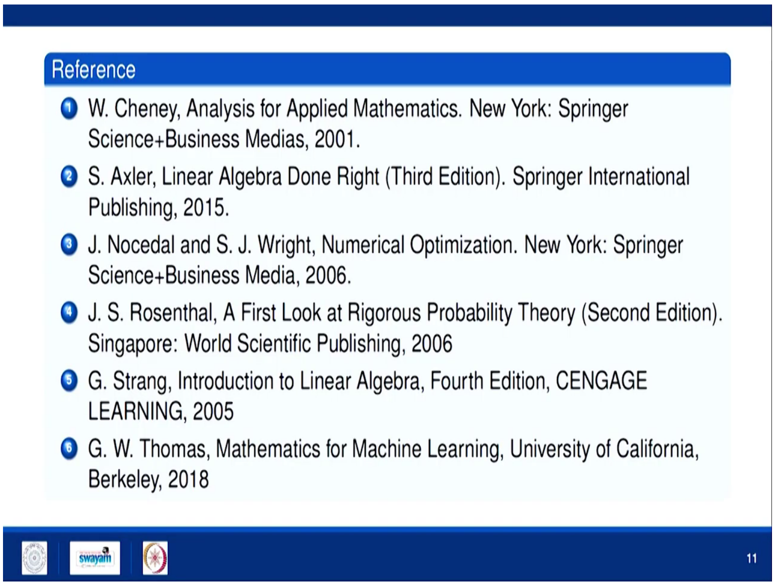
They also preserve L 2 norm that is the L 2 norm of Q X equals to L 2 norm of X. What it is saying this implies that multiplication by an orthogonal matrix can be considered as a transformation that preserve length because length of Q X equals to length of X, but may rotate or reflect the vector about origin that is obvious because they are we have seen examples of rotation matrices h the orthogonal matrix.

(Refer Slide Time: 32:31)



Another important properties of orthogonal matrices are the eigenvalues of orthogonal matrix have absolute value as 1. If Q is orthogonal, then determinate of Q equals to 1. If Q is an orthogonal matrix, then Q square equals to I because Q transpose Q equals to I and Q transpose is Q inverse only. If Q is orthogonal matrix, then nullity of Q that is the solution space of Q times X equals to 0 is 0 that is you do not have non-zero solution in the non-zero vectors in the null space of a orthogonal transformation.

(Refer Slide Time: 33:18)



So, in this lecture, we have told you about eigenvalues and eigenvectors. And then later part of this lecture, we have seen the properties related to eigenvalues and eigenvectors of two special types of matrices that is symmetric and orthogonal. I hope you have enjoyed this lecture. These are the references.

Thank you very much.