Essential Mathematics for Machine Learning

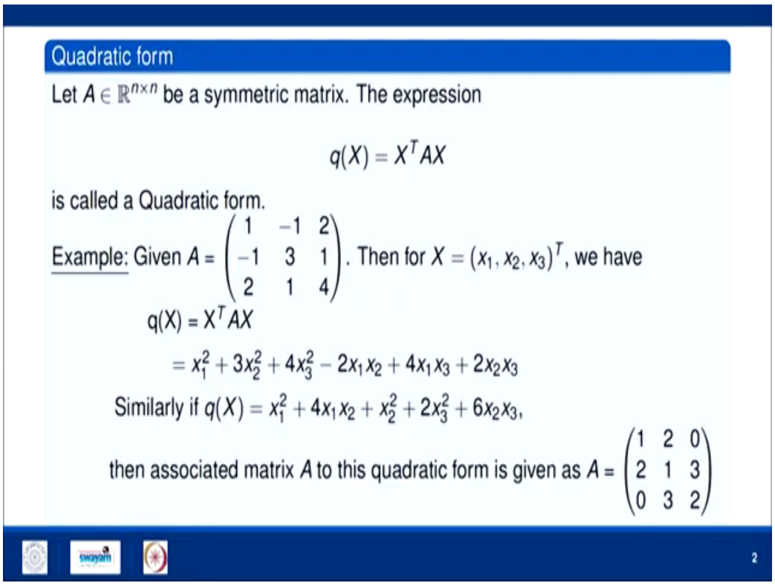
Prof. Sanjeev Kumar
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture - 09
Special Matrices and Properties

Hello friends. So, welcome to the module 9 of this course Essential Mathematics for Machine Learning. So, in previous lecture, I have told you about eigenvalues and eigenvectors and then, we have seen some properties of eigenvalues and eigenvectors, especially related to the eigenvalues and eigenvectors of symmetric and orthogonal matrices. So, we will continue that trend and in this lecture, we will see some more properties related to a special matrices.

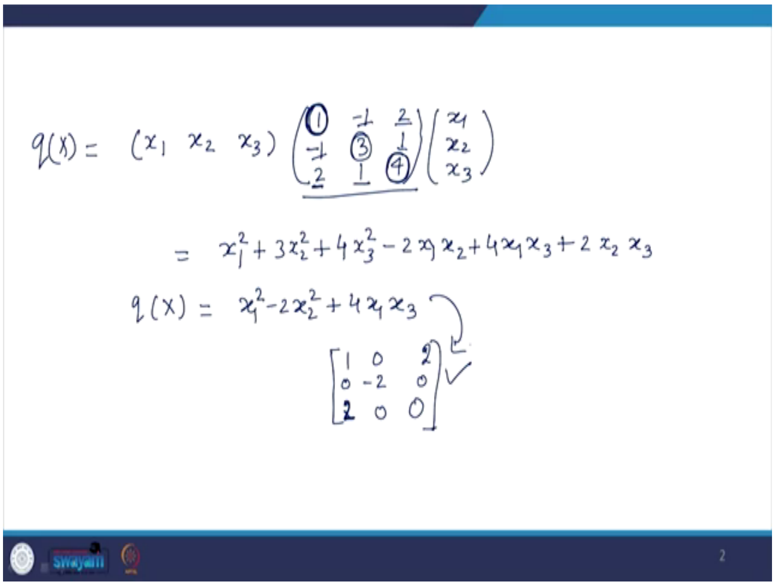
In particular, we will discuss about positive definite matrices, those are really important in convex optimization and hence, in machine learning. Because ultimately in machine learning you have to optimize the objective function or decision function whatever you are forming.

(Refer Slide Time: 01:28)



So, we will start this lecture with the definition of quadratic form. So, this is the definition of quadratic form. So, let A be a n by n real matrix which is symmetric also and capital X is a n-dimensional vector. Then, the expression q X equals to X transpose AX is said to be the quadratic form associated with matrix A. So, for example, you are having this particular matrix. So, it is a 3 by 3 matrix which is symmetric, then the quadratic form of A is given by q X, where capital X is a vector in R 3 that is having components x 1, x 2, x 3 and it is X transpose AX.

(Refer Slide Time: 02:24)



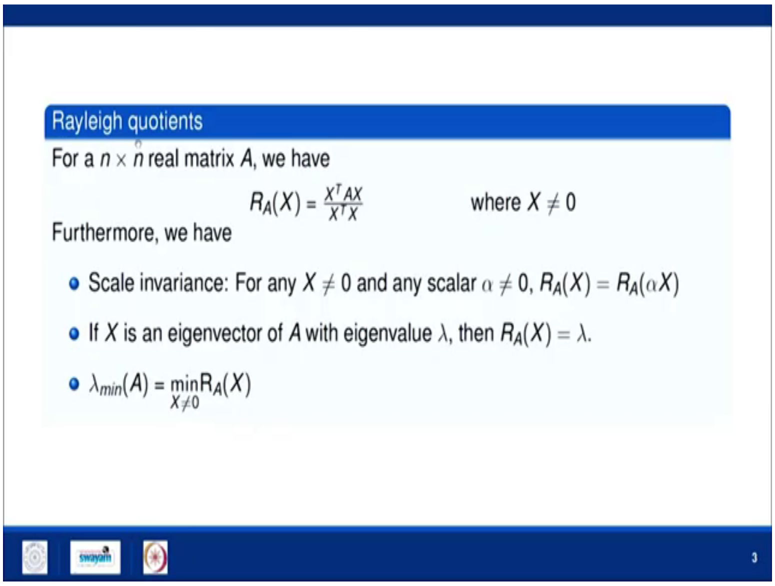
So, basically it will become X transpose will be something like this x 1, x 2, x 3 and then, it will be 1 minus 1 2, minus 1 3 1, 2 1 4. So, this is my matrix A and then, x 1, x 2, x 3. So, this equals to if you multiply this it comes out to be x 1 square. So, the coefficient of x 1 square will be this a 11 element of the matrix A plus similarly another diagonal element a 22 will be the coefficient of x 2 square. So, 3 x 2 square plus this third diagonal element that is dia 33 will be the coefficient of x 3. So, four x 3 square.

Now, what will be the coefficient of x 1, x 2? That sum of these two entries. So, minus 2 x 1 x 2. Similarly, the coefficient of x 1, x 3 will be sum of a 13 x a 31. So, plus 4 x 1 x 3 plus a 23 and a 32 sum of these 2 will be the coefficient of x 2, x 3 that is 2 times x 2, x 3. So, this is the quadratic form associated with matrix A.

Similarly, if a matrix quadratic form is given to you, let us you say I am having a quadratic form q x equals to some x 1 square minus 2 x 2 square plus 4 x 1 x 3, then what will be the corresponding matrix? So, matrix will be the coefficient of x 1 square will be the diagonal element a 11, then coefficient of x 2 square will be diagonal element a 22 that is minus 2 here, then x 3 coefficient of x 3 square is 0. So, I will write 0. So, these will be the diagonal elements.

Now, x 1, x 3 means related to these 2 entries a 13 and a 31. So, this is the sum of these two end the matrix A is a symmetric matrix. So, 22 will come here and rest of the element are 0 because their coefficients are 0. So, this is the matrix corresponding to this quadratic form. So, if a matrix is given, you can easily find out the associative quadratic form and vice versa, means if quadratic form is given you can easily find out the symmetric matrix.

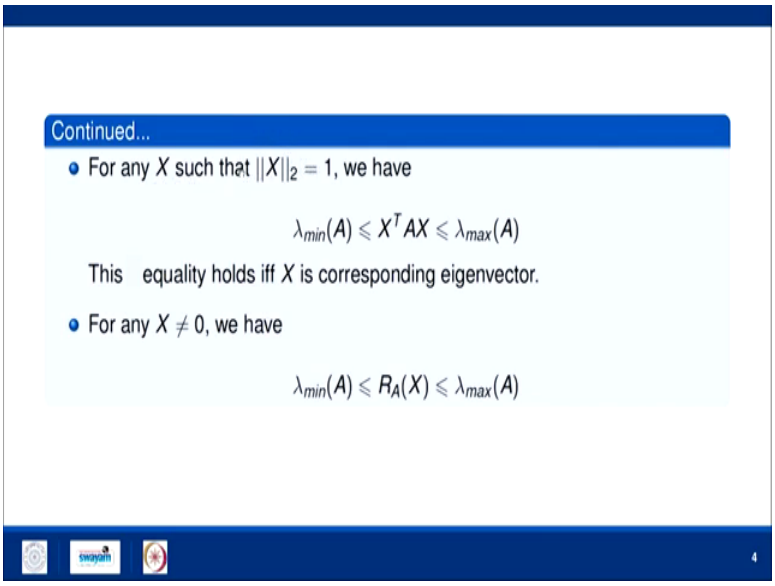
(Refer Slide Time: 05:57)



Another important concept is Rayleigh quotients. So, for a n cross n real matrix A, we have Rayleigh quotient of a each X transpose AX divided by X transpose X, where X is a non-zero vector of size n. Furthermore, we have scale invariance property of this Rayleigh quotient that is for any X not equals to 0 and any scalar alpha not equals to 0, Rayleigh quotient of X equals to Rayleigh quotient of alpha X.

Now, if X is an eigen vector of a with eigen value lambda, then Rayleigh quotient of these eigen vector X equals to the corresponding eigen value lambda. If you have to find out the minimum over all nonzero vector X and in the space of Rayleigh quotient, the minimum value will be the minimum eigen value.

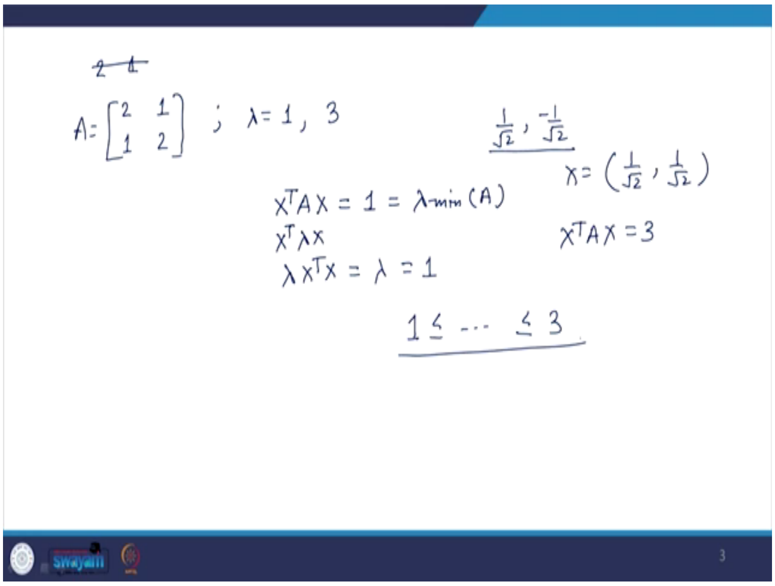
(Refer Slide Time: 07:10)



If you are having a unit vector X of n dimension, we have this inequality that the minimum eigen value of a less than equals to X transpose AX; because if I have divided by X transpose

X that is the length is 1. So, I have not written that denominator here, less than equals to maximum eigen value of A. This equality holds if and only if, X is corresponding eigen vector.

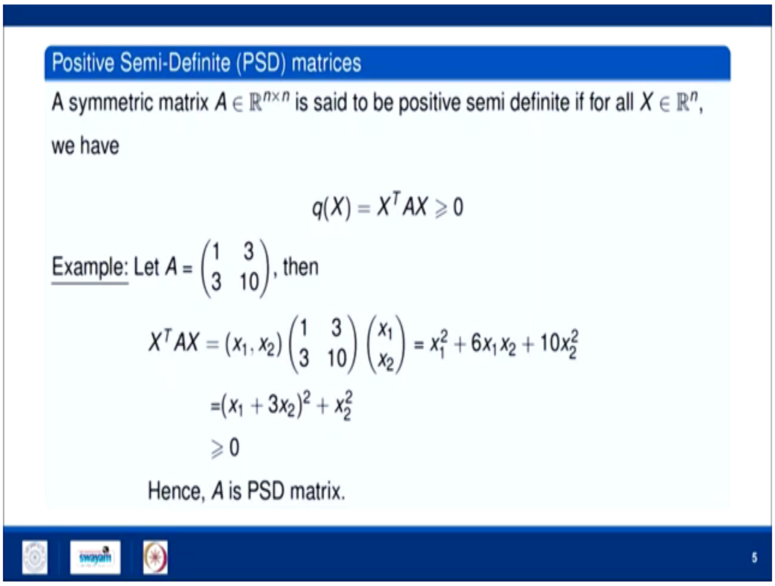
(Refer Slide Time: 08:00)



And this you can easily verify with the example which we have taken in previous lecture that is you are having 2 1; the matrix like A equals to 2 1, 1 2 eigen values are lambda equals to 1 and 3 and then, you take any vector X of unit length. So, let us I am taking 1 by root 2 and minus 1 by root 2, that is basically the eigen vector corresponding to lambda equals to 1. Now, if you find X transpose AX, this will come out to be 1, which is minimum eigen value of A and it is obvious because if I am taking this X here, then AX equals to lambda X.

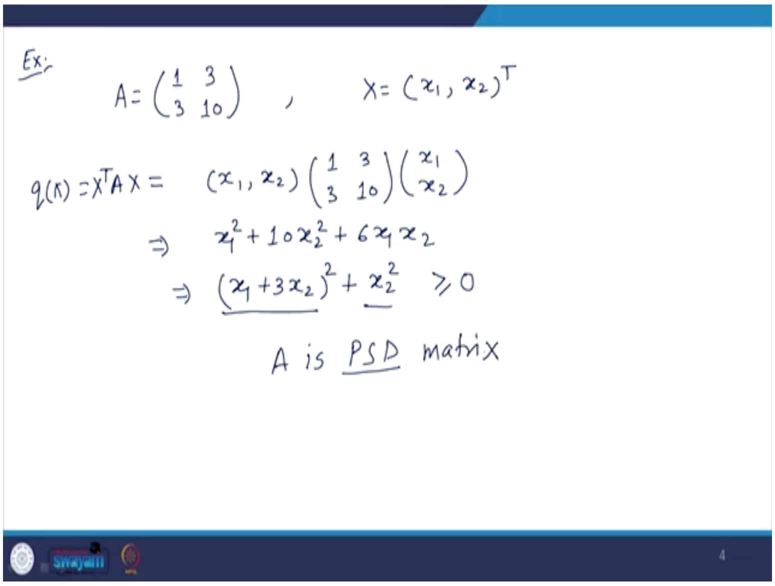
So, X transpose lambda X or lambda X transpose X, where lambda is the corresponding eigen value to X, eigen vector X and this is equals to lambda which is 1. Similarly, if I take this as 1 by root 2 and 1 by root 2 X, then X transpose AX will become 3, which is the bigger eigen value of A. If you take any other X, then the value will this value X transpose AX will lie between 1 and 3. It will not go out beyond 1 and 3 for all nonzero vectors X.

(Refer Slide Time: 09:47)



Now, we are coming to very important definition that is definition of positive semi definite matrices. So, a symmetric matrix A of size n by n and having real entries is said to be positive semi definite, if for all X belongs to Rn means for all real n-dimensional vectors, we have the quadratic form of A, that is X transpose AX is nonnegative. If this condition hold, we say that the matrix A is positive semi definite. So, for example, see this particular example. So, here A is 1 3, 3 10.

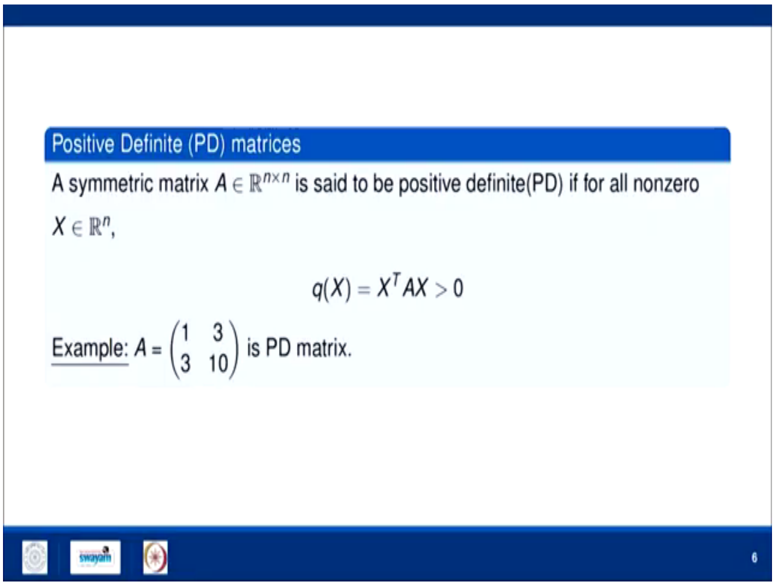
(Refer Slide Time: 10:42)



So, A is 1 3, 3 10. Now, take X equals to x 1, x 2 transpose because it is a 2 by 2 matrix. So, we have to take vector from R 2. Now, quadratic form q X is X transpose AX which is x 1, x 2 multiplied with 1 3, 3 10 and then, x 1, x 2. This will give you x 1 square plus 10 x 2 square plus 6 x 1 x 2. This I can write x 1 plus 3 x 2 whole square. So, it will become x 1 square plus 9 x 2 square plus 6 x 1 x 2 plus x 2 square and this is strictly positive, because why? Because it is sum of the squares.

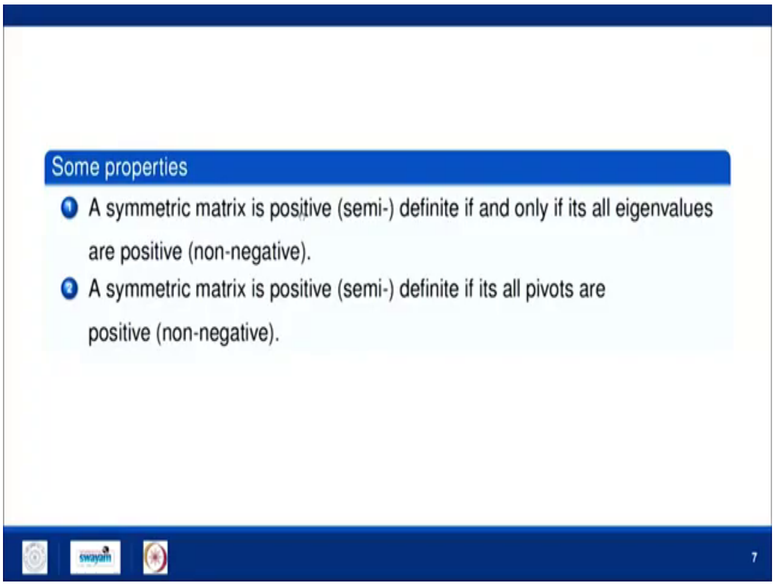
So, these square values can be at most 0 and it is sum of the values will be having the minimum value 0. So, obviously, it will be greater than equals to 0. Hence, A is a positive semi definite. So, in short, I am writing positive semi definite as PSD; P stands for positive, S for semi and D for definite; so, Positive Semi Definite matrix.

(Refer Slide Time: 12:46)



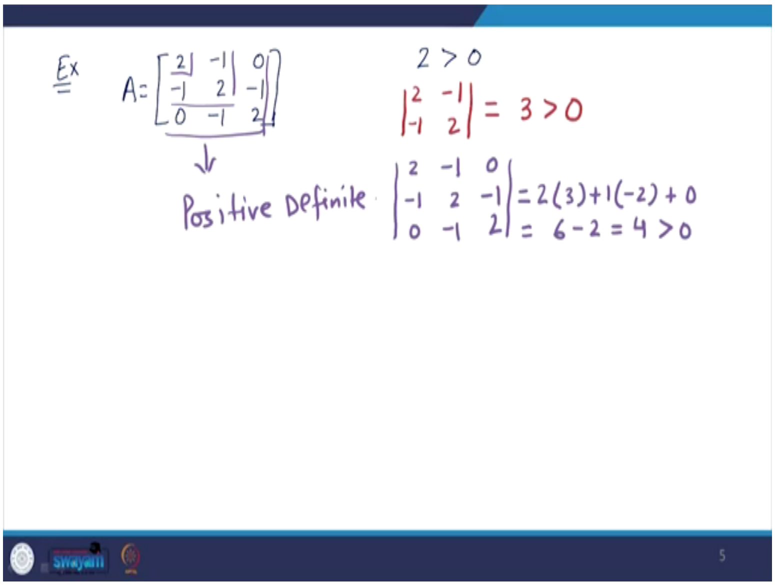
If the quadratic form X transpose AX for a symmetric matrix A is strictly positive, then we say the matrix is positive definite. So, if it is greater than equals to 0, the matrix is called positive semi definite; but if it is strictly positive, then we say the matrix is positive definite.

(Refer Slide Time: 13:16)



Now, some other checks for evaluating whether a given matrix is positive definite or not like this kind of thing. So, a symmetric matrix is positive semi definite or positive definite if and only if, it is all eigen values are non negative or a strictly positive. So, if all the eigen values of a matrix is greater than equals to 0, then the matrix is positive semi definite. If all the eigen values are strictly positive, then the matrix is positive definite. Another check is a symmetric matrix is positive or positive semi definite, if it is all pivots are positive or non-negative.

(Refer Slide Time: 14:21)

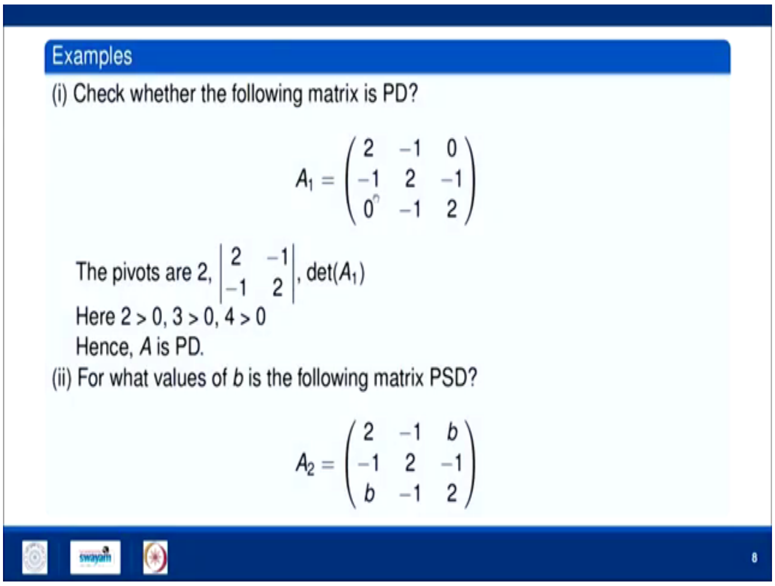


So, what is the meaning of this pivot kind of thing? So, just check this matrix. 2 minus 1 0, minus 1 2 minus 1 and 0 minus 1 2. One check is you can write the quadratic form and then, try to write the quadratic form as the sum of the squares. But that is not easy task for bigger matrices. So, pivot test is quite helpful in this case.

So, what is pivot test? Means all pivot, how I will check? First I will see this pivot. So, here it is 2, 1 by 1 pivot and 2 is greater than 0. Now, what I will do? Now, I will take this 2 by 2 pivot. So, it will become 2, determinant of 2 minus 1 minus 1 2. This comes out to be 4 minus 1, 3 which is again positive.

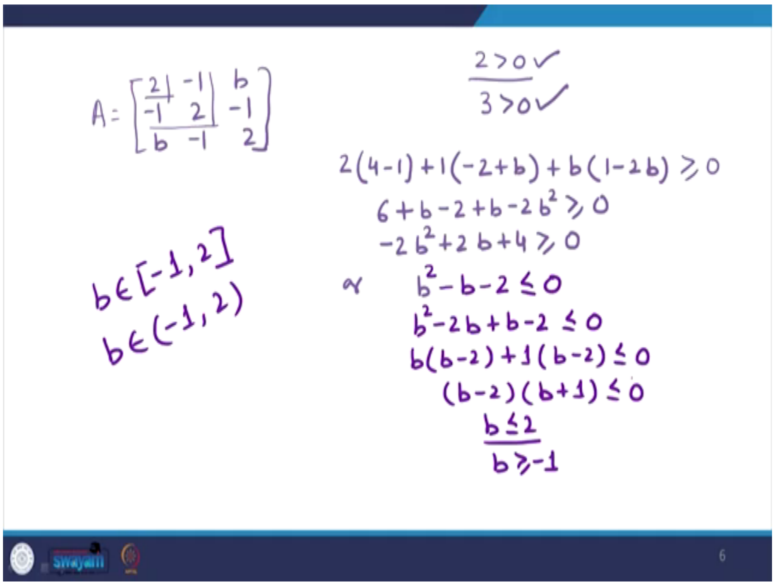
Then, I will take 3 by 3 pivot. So, 3 by 3 pivot is nothing in this case, it is the determinant of this matrix minus 1 2 minus 1, 0 minus 1 2. So, it is 2 times 3 plus 1 minus 2 plus 0. So, it is 6 minus 2 which is 4 which is again 0. So, all these pivots means this one, this one and the whole one are positive; even they are strictly positive. So, this matrix is a positive definite. If they are nonnegative means some of them are 0, then the matrix is called positive semi definite.

(Refer Slide Time: 16:27)



We are having one more very interesting example here. For what value of b is the following matrix positive semi definite?

(Refer Slide Time: 16:41)

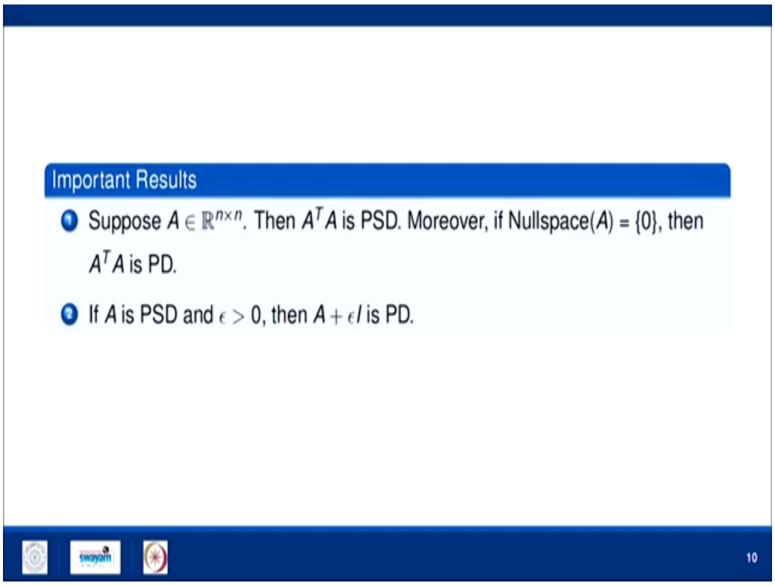


So, what is the matrix? Matrix is A equals to 2 minus 1 b, minus 1 2 minus 1, b minus 1 2. So, for what value of b the matrix is positive definite; positive semi definite. So, if you check this pivot the 1 by 1, it is 2. So, no issue. If you check another pivot, it is again 3, which is again positive. So, no problem. Now, what we need to do? We need to see the determinant of it. So, determinant will be 2 times 4 minus 1 plus 1 times minus 2 plus b plus b times 1 minus 2 b and for being the positive semi definite, it should be greater than equals to 0.

So, let us simplify it. So, 6 plus b minus 2 plus b minus 2 b square is greater than equals to 0 or minus 2 b square plus 2 b plus 4 is greater than equals to 0 or b square minus b minus 2 is less than equals to 0 because we have multiply by minus 1 both sides. So, inequality will change. So, it is b square minus 2 b plus b minus 2 less than equals to 0. So, b minus 2 plus 1 b minus 2 less than equals to 0. So, b minus 2 and b plus 1 less than equals to 0.

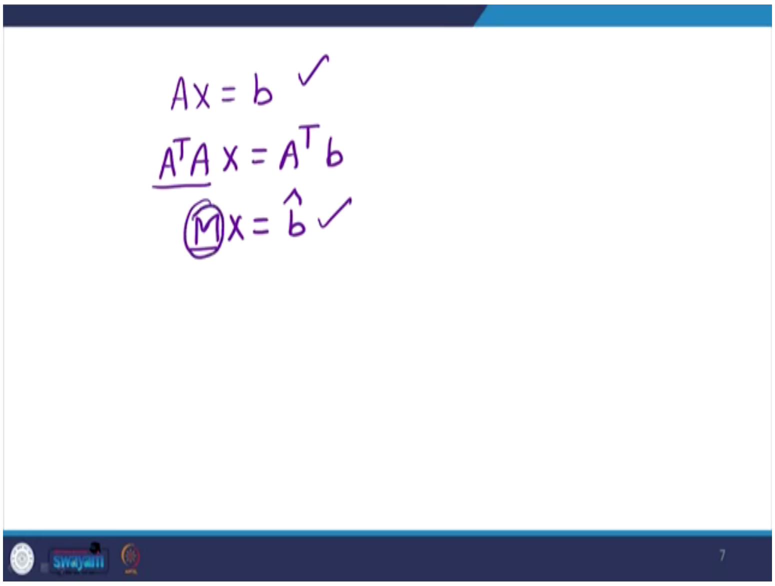
So, if b minus 2 less than equals to 0, we will be having b less than equals to 2 and if b plus 1 is less than equal to 0, we will be having b greater than minus 1. So, what I am having for all values of b, belongs to minus 1 to 2, the matrix A is positive semi definite. If you exclude these boundaries means minus 1 and 2, then it will become strictly positive and now, matrix will become positive definite.

(Refer Slide Time: 19:57)



Another very important result related to machine learning of this particular concept of positive definiteness is suppose A is a n by n matrix, any matrix real matrix which need not be symmetric. Then, A transpose A is positive semi definite. So, if you want in some algorithm positive definiteness on a certain matrix M, what you do? You multiply both sides by m transpose. So, now, it will become a new system.

(Refer Slide Time: 20:41)

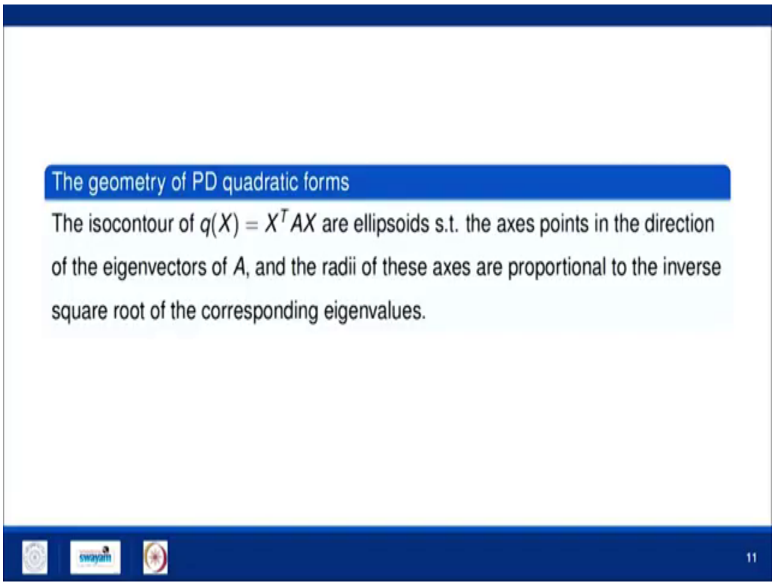


So, for example, you are having some algorithm, where you have to solve AX equals to b and this algorithm requires that s would be positive semi definite. So, what you do you multiplied by A transpose both side, A transpose b and now, write this system as MX equals to some b k, where M is A transpose A which is positive semi definite and b is just A transpose b.

So, the solution of this will be the solution of this one because both are equivalent system and now, your coefficient matrix is positive semi definite. Moreover, if null space of A is having only 0 vector, then A transpose A is positive definite. Then, they that you can easily observe. If null space of A is having only 0 vector; means A is full rank matrix because nullity is 0. So, rank will be full and if rank is full, then none of the eigen value will be 0, means all the eigen values are strictly positive.

So, hence, it will be positive definite. Another important aspect is if A is a positive semi definite and epsilon is some a constant which is positive, then A plus epsilon I is positive definite. So, whatever is small epsilon you take, it if it is positive and you are adding this epsilon in the diagonal elements of A, then it will make it positive definite because if any eigen value will be 0. Due to this epsilon, it will become a strictly positive.

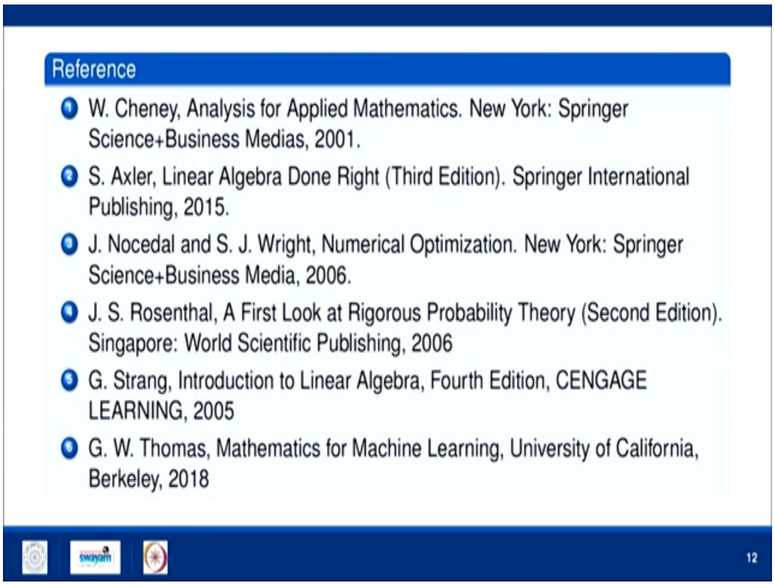
(Refer Slide Time: 22:38)



So, if we talk about the geometry of quadratic form, both especially the positive definite quadratic form the isocontour of this quadratic form X transpose AX which is a polynomial of degree 2 are ellipsoid. So, that polynomial of degree 2 that quadratic equation will give you ellipsoid, in such a way that the axes points in the direction of eigenvectors of matrix A.

So, again, eigenvector is giving the direction and the radii of these axes are proportional to the inverse square root of the corresponding eigen values. If you want to see this concept further, because I have not given proof here.

(Refer Slide Time: 23:27)



So, what you can see? You can see the go to reference number 6 and there you will find the proof of these. So, in this lecture, we talked about quadratic form and then, the concept of positive definiteness. This concept, we will use in many of the subsequent lectures. I hope you enjoyed this lecture.

Thank you very much.