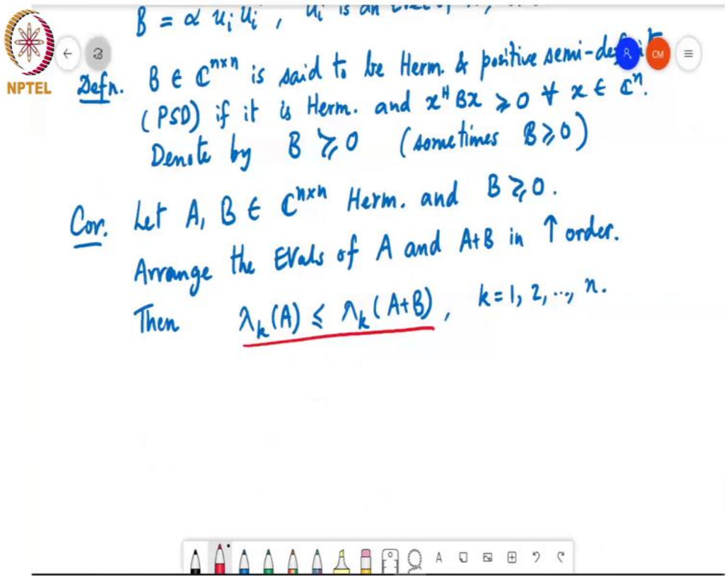
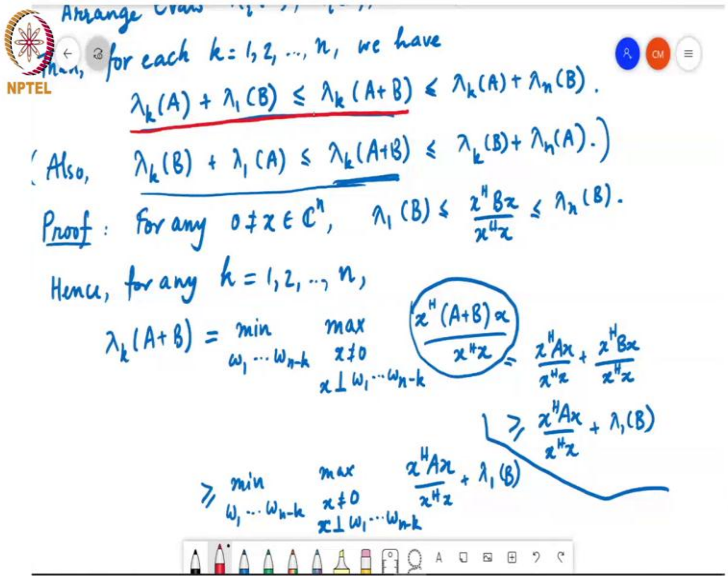
Matrix Theory
Professor Chandra R Murthy
Department of Electrical Communication Engineering
Indian Institute of Science Bangalore
Positive semi-definite matrix, monotonicity theorem
and
interlacing theorems

(Refer Slide Time: 00:13)





So, now in order to proceed further, so I need one more definition, it is just a small extension of what we already know. So, B in C to the n cross n is said to be Hermitian and positive semi-definite, which I will often abbreviate by PSD, it is Hermitian, and x Hermitian Bx is greater than or equal to 0 for every x in C to the n. So, the only difference from a positive definite matrix is that the inequality becomes a greater than or equal to 0.

And it is okay, for it to hold for every x in C to the n, we do not have the restriction that x must be a non-zero vector. Of course, for defining positive definite matrices, I cannot allow x equal to 0 because x Hermitian Bx will be equal to 0, so I would not be able to satisfy a strict inequality constraint if I want to define positive definite matrices, but for positive semi-definite matrices, we can allow x to be equal to 0 because after all, the condition is that it should be greater than or equal to 0.

And we will denote this by B greater than or equal to 0 like this and I may even write it as B greater than or equal to 0. But unless, unless I explicitly say otherwise, B greater than or equal to 0 means positive semi-definite, it does not mean that the entries, all the entries of B are real and non-negative.

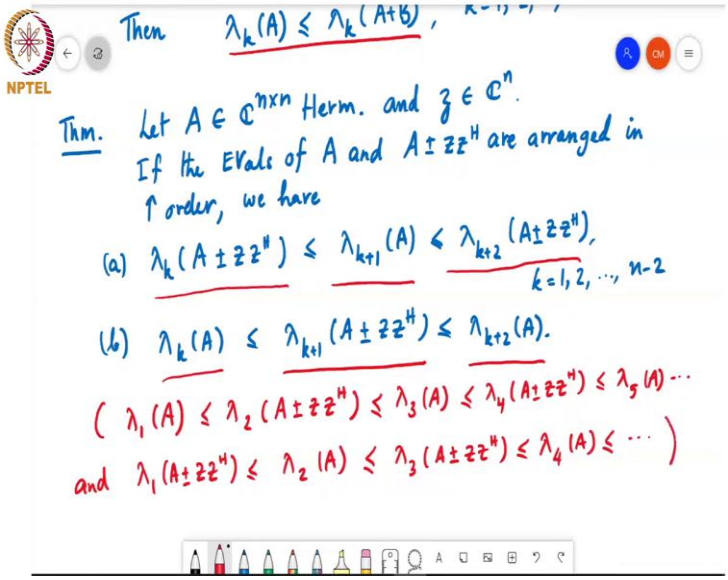
So, once we define these positive semi-definite matrices, we can now talk about what is called a monotonicity theorem, which says that all the eigenvalues of a Hermitian matrix cannot, none of the eigenvalues can decrease, or they will all increase but either increase or remain the same, if you add a positive semi-definite matrix to the matrix, so it is the following corollary.

Let A and B be n cross n Hermitian symmetric matrices and B is positive semi-definite arrange the eigenvalues of A and A plus B in increasing order. Then lambda k of A is less than or equal to lambda k of A plus B, 1, 2 up to n, this directly follows from the previous theorem because, in fact, all you need is the lower inequality.

One immediate consequence of x Hermitian Bx being greater than or equal to 0 is that lambda, or (())(04:31) here, this lower inequality, I will use a different color, this lower inequality because one immediate consequence of lambda, of B being a positive semi-definite matrix is that x Hermitian Bx is greater than or equal to 0 for all x and so, all the eigenvalues are greater than or equal to 0. In particular, lambda 1 of B is greater than or equal to 0.

And so, lambda k of A plus B is to be greater than or equal to lambda k of A plus some non-negative quantity. So, if I drop this non-negative quantity, the inequality will still hold. And so, lambda k of a will be less than or equal to lambda k of A plus B. So that is what the result also says.

(Refer Slide Time: 5:23)



Now, another thing is that once you start looking at positive semi-definite matrices, you can say more about these eigenvalues. It is not just the smallest and largest eigenvalues of B that you need to use, in order to bound the kth eigenvalue of A plus B, you can prove much more sophisticated results about the location of these eigenvalues. And these kinds of bounds often take the form of what is known as an interlacing theorem. And there are many interlacing theorems in linear algebra.

So, we will discuss a few of these interlacing theorems. The 1st interlacing theorem we are going to discuss is for the case where B is a rank 1 matrix. So that is the following theorem. So, as before let A be an n cross n Hermitian symmetric matrix and z be a vector in C to the n, if the eigenvalues of A and A plus or minus zz Hermitian, this z is the same as this z, as I wrote it differently, are arranged, this z plus z, z A, sorry A plus zz Hermitian and A minus zz Hermitian are two different matrices.

But whatever I am going to say is valid for both those matrices. So, that is why I am combining them and writing it as A plus or minus zz Hermitian in increasing order, we have two parts to it lambda k of A plus or minus zz Hermitian is less than or equal to lambda k plus 1 of A is less than or equal to lambda k plus 2 of A plus, plus or minus zz Hermitian.

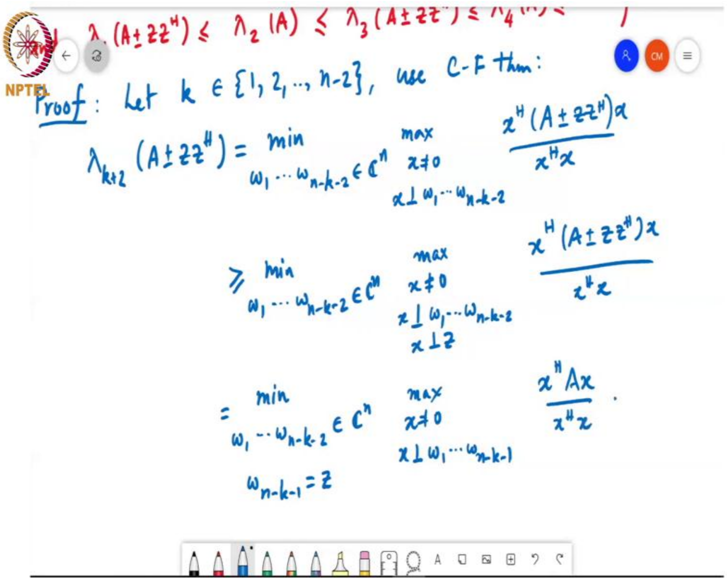
And this is true for k equal to 1, 2 up to n minus 2. Of course, this one I have k plus 2 here, k can only go up to n minus 2 and B lambda k of A is less than or equal to lambda k plus 1 of A plus or minus zz Hermitian is less than or equal to lambda k plus 2. So, what this result is saying is that the k plus 1th eigenvalue of A can be bounded by, it can be bounded below by the kth eigenvalue of A plus or minus zz Hermitian and upper bounded by the k plus 2th eigenvalue of A plus or minus zz Hermitian.

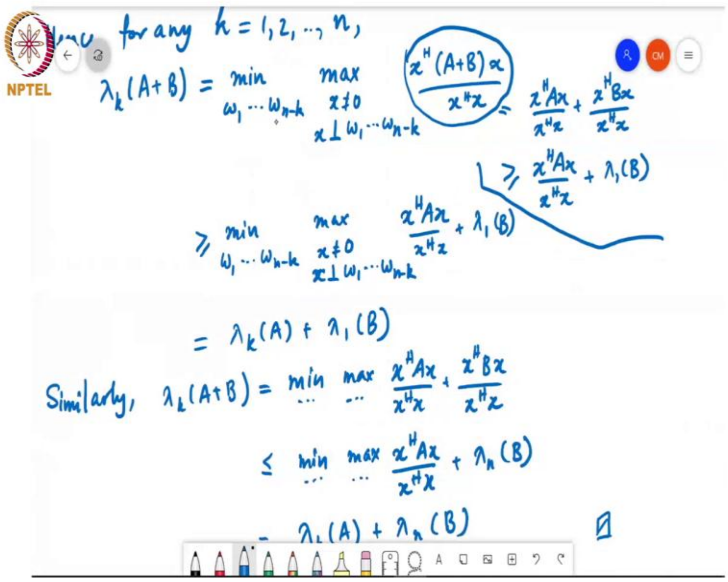
And similarly, the k plus 1th eigenvalue value of A plus or minus zz Hermitian is lower bounded by lambda k of A and upper bounded by lambda k plus 2 of A. So, in particular if I write all the eigenvalues from k equal to 1, 2 up to n minus 2 of A and A plus or minus zz Hermitian then I approximately this is what I have lambda 1 of A that is this when I put k equal to 1 here, is less than or equal to lambda 2 of a plus or minus zz Hermitian.

So, it is less than or equal to, so, this is lambda 2 and k plus 2 becomes lambda 3 of A less than or equal to. Now, lambda 3 of A, I can take A equal to 3 here, lambda 4 of A plus or minus zz Hermitian less than or equal to lambda 5 of A, and so on.

And if I start with this part A here, lambda 1 of A plus or minus, zz Hermitian, less than or equal to lambda 2 of A, less than or equal to lambda 3 of A plus or minus zz Hermitian less than or equal to lambda 4 of A, less than or equal to etcetera. So, that is what these two inequalities say. Let us see how to show this.

(Refer Slide Time: 11:24)





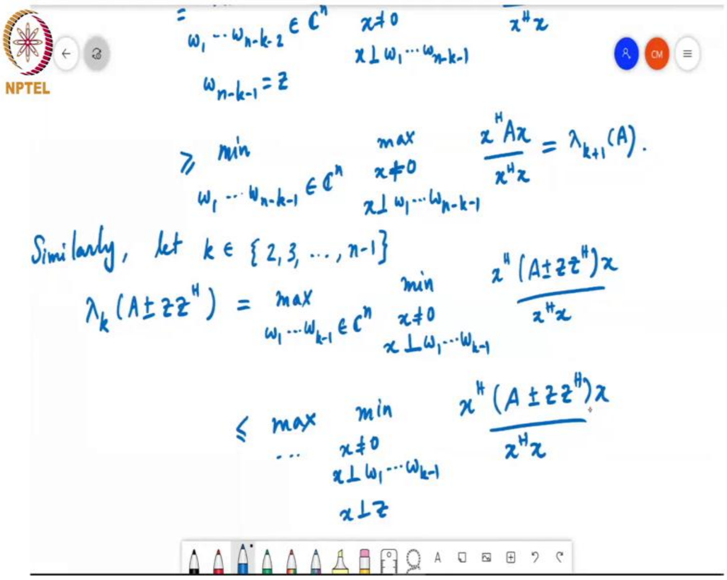
So, the starting point is again, the Courant-Fischer theorem. So, let k be some number n minus 2, that is what we have in the condition here and use the Courant-Fischer theorem. So, I will start with lambda, this quantity lambda k plus 2, A plus or minus zz Hermitian. This is equal to the min over w1 up to w, can go back here, so for lambda k, it is a min over w1 through wk.

So, if I want lambda k plus 2, I must do min w1 to wn minus k minus 2, these in C to the n, then max over x not equal to 0 and x perpendicular to w1 through wn minus k minus 2, x Hermitian A plus or minus zz Hermitian x over and this in turn is greater than or equal to. So, this is a slightly clever step in the proof the min over the same quantity wn minus k minus 2 in C to the n, the max over x not equal to 0 and x perpendicular to w1 through wn minus k minus 2 and x perpendicular to z, x Hermitian A plus or minus zz Hermitian x over x Hermitian x.

Why is this true? It is true because all I am doing is adding one more constraint here. So, whatever max this could have achieve, maybe this cost function cannot achieve as, as great maximum value, because it is a more constrained optimization problem. So, this will be smaller than this.

Now, if x is perpendicular to z, this term here is x Hermitian z, z Hermitian x. So that term just drops off. So, I can drop that term and write it in the following way. This is equal to the min over w1 through wn minus k minus 2 in C to the n. And I will just define a wn minus k minus 1 to be equal to z. And then I will say the maximum is over all x not equal to 0, and x being perpendicular to w1 up to wn minus k minus 1, that is the same as saying x is perpendicular to z. And my cost function is now x Hermitian Ax over x Hermitian x.

(Refer Slide Time: 15:21)



Now, instead of minimizing it subject to this constraint that wn minus k minus 1 equals z, if I just drop this constraint, and allow wn minus k minus 1 to be any vector in C to the n, I can possibly achieve a lower minimum than what is achieved by this cost function here of this objective function here, rather this optimization problem here.

So, I have a further lower bound by allowing wn minus k minus 1 to be anything. So that is the min over w1 through wn minus k minus 1 in C to the n, the maximum x not equal to 0, x perpendicular to w1 through wn minus k minus 1 of x Hermitian Ax over x Hermitian x. And by Courant-Fischer theorem, this is exactly what we, equal to lambda k plus 1of A.

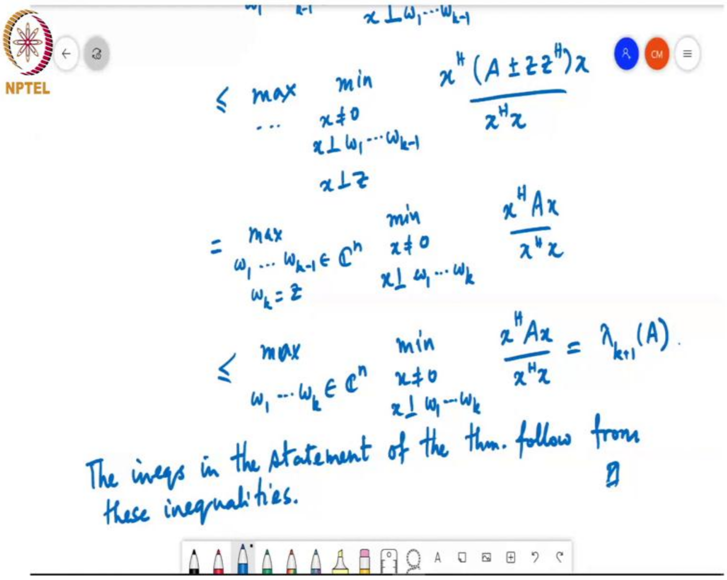
So, you can see that the proof involves this interesting step of saying, so, if x was perpendicular to Z, you would get a lower bound on the 1st cost function, and then you push the constraint into this wn minus k minus 1, then you allow that to become arbitrary and both those steps are lower dotting steps. And that gives you lambda k plus 1 of A.

Similarly, I can do the other way. So, let k be in, so the theorem, proving the theorem involves proving 4 inequalities, and this proves one of them. I will do one more and but, from these inequalities, the theorem follows 2, 3 up to n minus 1. So, if I look at lambda k of A plus or minus zz Hermitian.

This is equal to, now I will use the max min version. This is the max over w1 to wk minus 1 in C to the n of the minimum over x not equal to 0, x perpendicular to w1 through wk minus 1, x Hermitian A plus or minus zz Hermitian x over x Hermitian x just Courant-Fischer theorem, which is then less than or equal to the max over the same constraint as the previous, the minimum over x not equal to 0, x perpendicular to w1 through wk minus 1.

And now, since I have thrown an extra constraint that x is perpendicular to z of x Hermitian A plus or minus zz Hermitian x over x Hermitian x. And now, since I have thrown in this extra constraint that x is perpendicular to z, this minimum may not be able to achieve as low a minimum as this optimization problem and therefore, this is an upper bound.

(Refer Slide Time: 19:25)

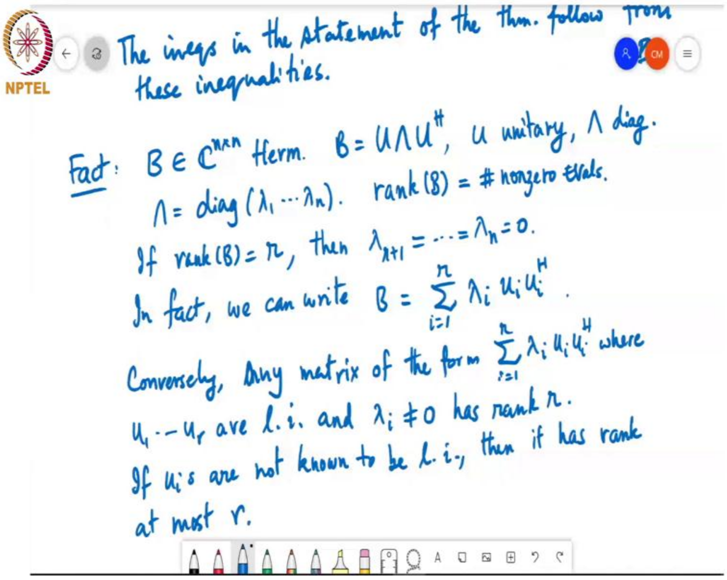


And this now the next steps are exactly the same as before. So, I will do the max over instead of I have w1 through wk minus 1 in C to the n, C to the n and I will set wk equal to z of the minimum x not equal to 0, x perpendicular for, to w1 through wk, x Hermitian and I can drop this term because x is perpendicular to z here. So, Ax over x Hermitian x and then I will neglect this constraint wk equal to z and thereby I can possibly achieve an even higher maximum then this optimization problem.

So, I have max over w1 through wk in C to the n the min over x not equal to 0, x perpendicular to w1 through w k of x Hermitian Ax over x Hermitian x and by Courant-Fischer theorem is exactly equal to lambda k plus 1 of A. So the inequalities in the statement of the theorem follow from these inequalities.

Notice that we use the max min formulation to prove an upper bound on, an eigenvalue and we use the min max formulation to prove a lower bound on the on an eigenvalue, and this is something that you will notice in many results that we show going forward.

(Refer Slide Time: 21:52)

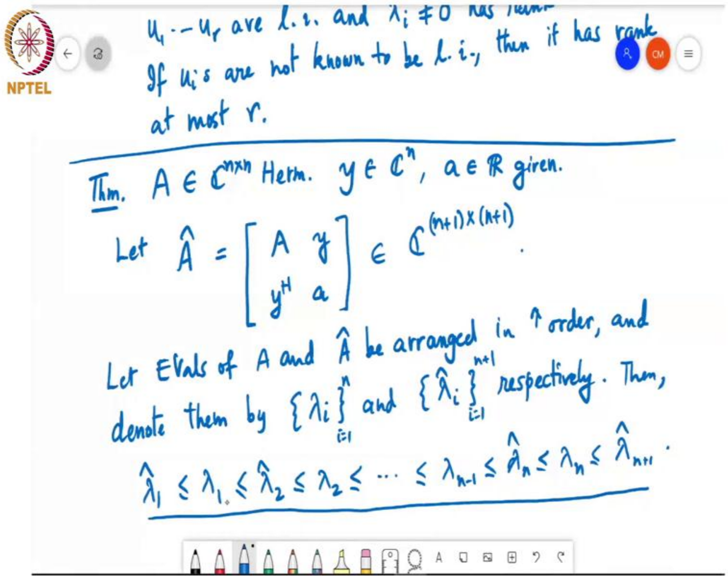


Now, one useful fact to keep in mind is that if, if B is in C to the n cross n, let me put it this way, and this is a Hermitian symmetric matrix then that means that it is unitarily diagonalizable and it is a non-defective matrix. And so, for such a matrix we can write B as u lambda u Hermitian where the matrix u contains the eigen vectors of B and lambda is a diagonal matrix containing the eigenvalues of B. So, u is unitary and lambda is diagonal. Now, so, let us say lambda is equal to, sorry, lambda 1 through lambda n. So, these lambda 1 to lambda n are the eigenvalues of B.

Now, in this case because B is non defective, the rank of B is equal to the number of non-zero eigenvalues and in particular, if the rank of B equals say r, then only some, only some are eigenvalues here we will be non-zero and the remaining will be equal to 0. So, I can say that lambda r plus 1 equal to etcetera up to equal to lambda n is equal to 0.

And so, in fact, we can write B as sigma i equal to 1 to r, lambda i times ui ui Hermitian and conversely any matrix of this form, of the form sigma i equal to 1 to r lambda i, ui ui Hermitian, where u1 up to ur are linearly independent has a rank at most r and if, if all these lambda i's are not equal to 0, let me put it this way, we will write it in the following way and lambda i not equal to 0 has rank r. If u i's are not known to be (linear), u i's not known to be linearly independent, then it has rank at most r.

(Refer Slide Time: 26:48)



This fact will turn out to be quite useful. So, for example, a rank 1 matrix can be written as some lambda times u u Hermitian for some vector u which is non-zero. So, a rank 1 Hermitian symmetric matrix can be written as some lambda 1 times u u Hermitian and so on.

So, the next result that I want to share is also an interlacing theorem and this is about what happens if you pair a matrix by a row and column to get a matrix whose size is one more than the matrix you started with. So, it reads like this. So, suppose A is in C to the n cross n and is Hermitian and y is in C to the n and A is a real number and these are given, then let A hat, denote the matrix A and y, y Hermitian and the small a, this is a matrix of size n plus 1 cross n plus 1, so question is, how are the eigenvalues of A related to that of A hat?

So, remember that A hat has n plus 1 eigenvalues and A has only n eigenvalues. So, let the eigenvalues of A and A hat be arranged in increasing order and denote them my lambda i and lambda hat i respectively. So, this is actually i equal to 1 to n and this is i equal to 1 to n plus 1.

Then lambda 1 hat is less than or equal to lambda 1 is less than or equal to lambda 2 hat is less than or equal to lambda 2, less than or equal to lambda n minus 1 less than or equal to lambda n hat less than or equal to lambda n is less than or equal to lambda n plus 1 hat. So, this is the interlacing theorem.

So, what it says is that the largest eigenvalue of this matrix is going to be bigger than the largest eigenvalue of A, the smallest eigenvalue of this matrix lambda 1 hat is going to be smaller than the smallest eigenvalue of A. So, all the eigenvalues and but not only that, all the eigenvalues of A hat interlace between pairs of eigenvalues of A. So, lambda 2 hat is going to be between lambda 1 and lambda 2, lambda 3 hat is going to be between lambda 2 lambda 3, lambda n hat is going to be between lambda n minus 1 and lambda n.

So, the last part, one largest or the 2nd largest eigenvalue of A hat is between the largest eigenvalue of A and the 2nd largest eigenvalue of A, but the largest eigenvalue of A hat is greater than or equal to the largest eigenvalue of A.