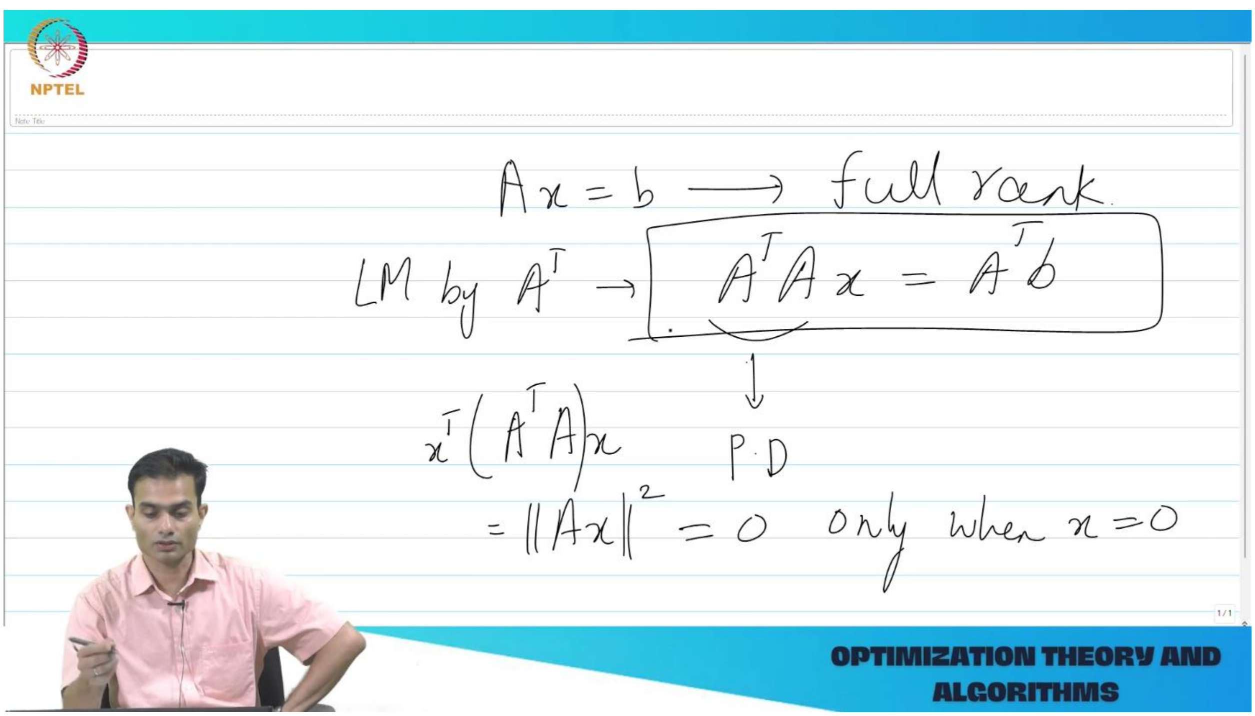
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 06
Lecture - 41

Discussion on doubts

Ah, how do you find the order of complexity for eigen decomposition? I said it was $O(n^3)$, ah I did not find the order of complexity I only stated it, okay. So, why is it $O(n^3)$? It is not something that I have proven here, it is something that you do in a linear algebra class, okay. It is a standard result, like Gram-Schmidt process. Ah, so we are not proving it, ah. In the affine space result of the expanding subspace theorem, is the x_0 only a general vector or a specific x_0? The x_0 that we have taken so far as the starting point of the CG method is completely general.

There is nothing special about it. So, you can take it to be whatever favorite point you want. What people—if you do not have any a priori knowledge about it—you would initialize it by a random variable. Or there is nothing special about it.

None of the results depend on it being something special. The only thing special is about the starting. Have we done that yet? The starting conjugate vector? No, right? We will come to that. Is the conjugate direction method and the conjugate gradient method valid only for solving $Ax = b$ when A is symmetric positive definite? It is a very small subset of problem. So, answer is: Is it only for solving $Ax = b$ when A is symmetric positive definite? Yes, no.



Should be? Answer is no, I think we did this in class. What if I give you an A which is not positive definite and I want to use the CG method, what do I do? Transformation? What transformation? At the back there, or do we have to go back to something like gradient descent? Anyone? Come on. If the eigenvalues are... But then you have changed the problem. If I change the signs of the eigenvalues, I make it... I change the problem itself. So, this is the problem statement. I want to do conjugate gradient method. You have seen that in whatever we have done, we have critically depended on the fact that A is positive definite symmetric positive definite. But now I give you A which is not.

What do we do? Any idea? I am surprised because we, I actually I derived this. I thought the last time, you know, in fact. No, that is in a quadratic form. The skew-symmetric part does not contribute, right. Pre-multiply by A^T. If I left-multiply by A^T, what do I get? $A^T A x = A^T b$. What is this $A^T A$ matrix? Is it symmetric? Is it positive definite? It is positive semi-definite. Why? Because it can have some eigenvalues which are 0, right.

As long as this is positive definite, right, you can see that $x^T A^T A x$—what is this? This is $\| Ax \|^2$, right. If A is full rank, then this norm can be 0 only when $x = 0$. In other words, for non-zero x, this cannot go to 0, right. So, only when $x = 0$.

So, to solve. So, as long as you tell me this is full rank, I can do this, right? I can solve this system of equations. The solution to this is the same as the solution to $x = b$. What was the problem with this? I mean, slight disadvantage of doing this? We are not inverting it. Remember, in solving in the conjugate gradient method, we can write formally. I can write the solution is equal to $A^{-1}b$, this is formally the solution, but do I ever evaluate A^{-1}? In whatever, we never evaluate A^{-1}.

So similarly, we will never evaluate $A^T A^{-1}$, it is just symbolically writing it as a solution, but we will have to do it because it is computationally very expensive to do it, okay. So similarly here, we want to do it, but what is the slight disadvantage of going this way? The condition number, right? The condition number of $A^T A$ is the square of the condition number of A, right. So, in the subsequent few lectures, we will see what role the condition number plays, okay. So, this is, but I mean if you are telling me that A is not symmetric, I have no option. I have to go through this, okay. So, there is some cost to be paid.

So, it is not. So, the conjugate gradient method is not very restrictive, it can solve any $Ax = b$, of course, it makes sense to solve $Ax = b$ only when A is full rank. So, it is not very restrictive, okay.