Course Name: Optimization Theory and Algorithms

Professor Name: Dr. Uday K. Khankhoje

Department Name: Electrical Engineering

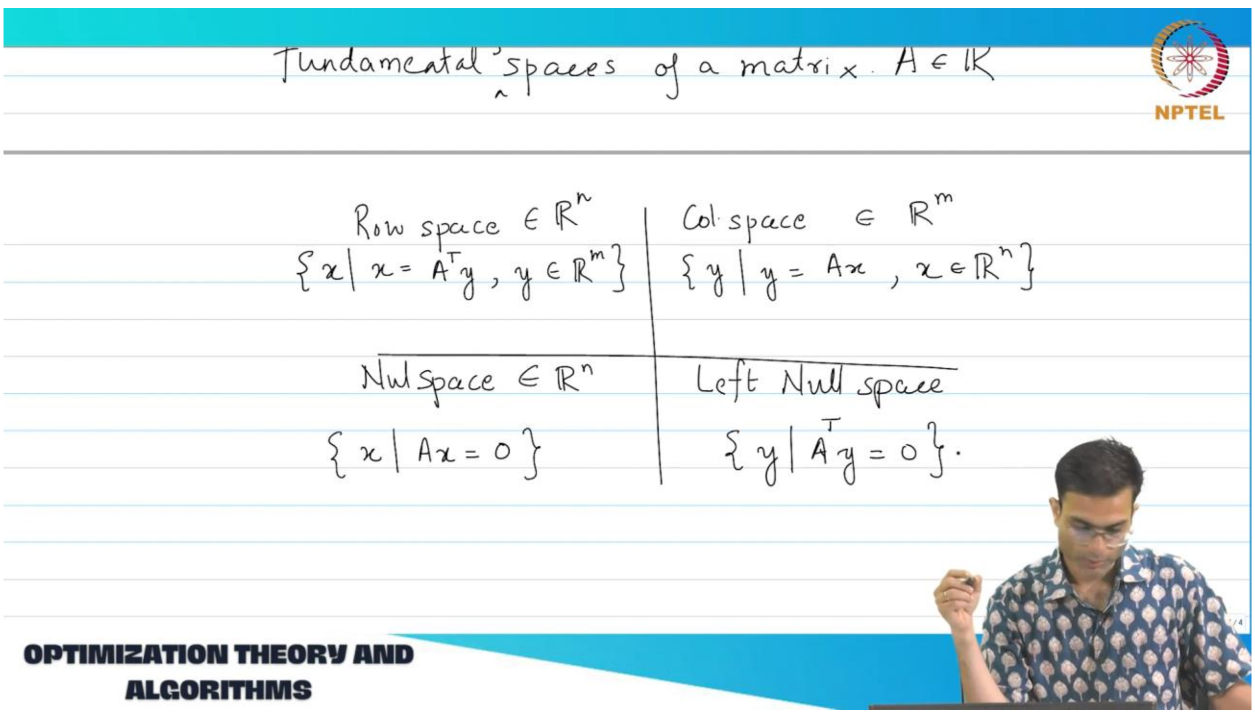
Institute Name: Indian Institute of Technology Madras

Week – 01

Lecture - 05

Summary of background material - Linear Algebra II

Alright, so, this is about norms, now we need to briefly touch upon the fundamental spaces corresponding to a matrix. So, those of you who have done linear algebra will recognize this very easily. ok. So, let us start with the most common subspace of a matrix. First, I would start with what is called the column space ok. So, I am going to split this like this.



In plain English, how would I define the column space of a matrix? The space spanned by columns. The space spanned by the columns of A. as simple as that. That means take all the columns of A and generate any linear combination of that, that is defined as the column space ok.

Is there a very compact way of writing this? Ax. right? So, I am going to define this as any vector y such that $y = Ax$. Is there any restriction on x? No right, $x \in \mathbb{R}^n$. ok. What is the dimension of this space? m right because $A \cdot x$ will become a vector of length m, right.

So, this is $\mathbb{R}^m$ right? it is not going to be exactly this is going to be a subset of $\mathbb{R}^m$ because we do not know if we can cover all the vectors in $\mathbb{R}^m$. It is not a given. ok. after we define the column space the next intuitive space to define is the null space right. How do I define the null space of a matrix A? Those vectors x such that $A \cdot x = 0$ right.

So, I am going to define this as x such that $A \cdot x = 0$. ok. So, obviously the null space every x is n dimensional. So, the null space is living inside $\mathbb{R}^n$ ok. Now, that we have got the column space and the null space there are you can see two missing corners over here.

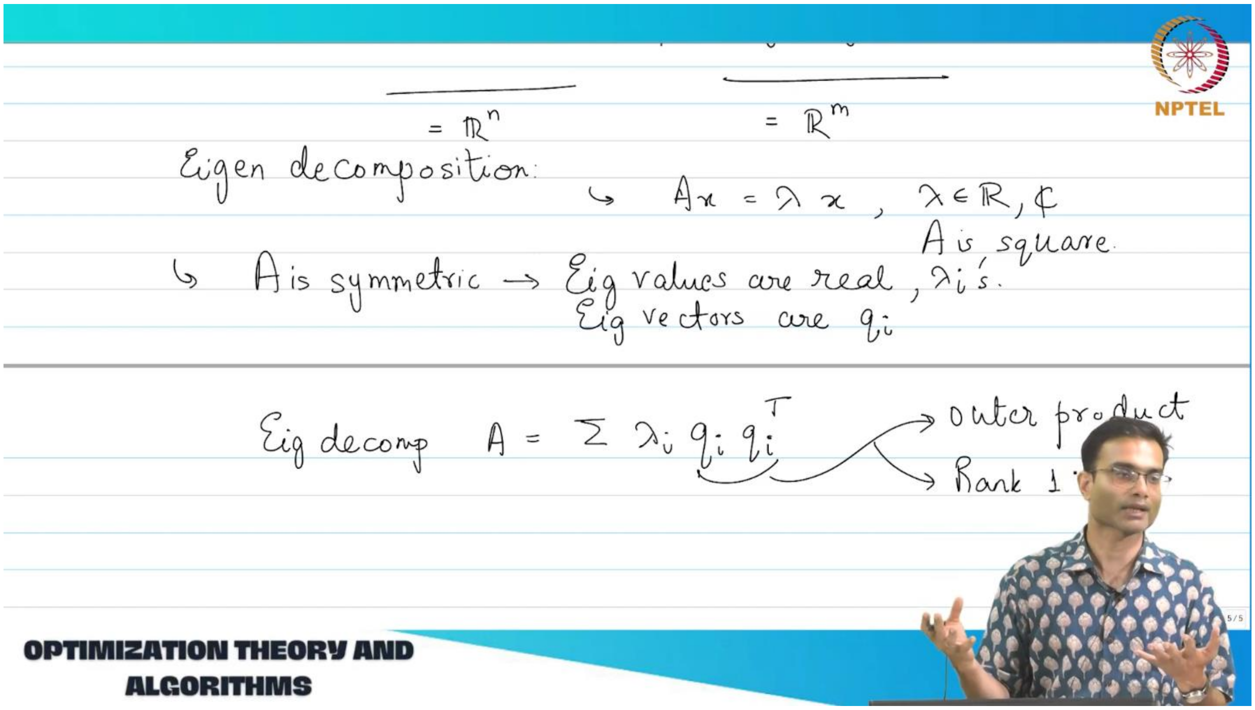
So, what should I write over here in the top left? Row space. Row space simply is the space spanned by the linear combination of the rows of A. Rows of A are the same as columns of $A^\top$ right. So, very analogously I can also define this as x such that $x = A^\top y$ ok and is there any restriction on y? No right.

So, $y \in \mathbb{R}^m$ ok and x also obviously lives in $\mathbb{R}^n$ ok. What is left? Left null space. How do I define the left null space? All those y's such that $A^\top y = 0$.

ok. So, these are called the fundamental subspaces of a matrix ok. Does there are two very surprising, but very elegant properties of these subspaces can anyone mention at least one of them? So, if I which which ok. So, one student has said that they are orthogonal which two are orthogonal? Column space and null space. No, I mean they are the wrong sizes. So, if I take the row space and I take the null space these are both vectors living in $\mathbb{R}^n$.

It turns out that these vectors are orthogonal to each other right. So, between this and this there is an orthogonality relation that is one thing between this and this there is an orthogonality relation ok. It is not very obvious, it is easy to prove this, but it is not very obvious just from the definition that they should hold to. So, this is one very very important property, there are algorithms optimization algorithms in signal processing which make use of this property to solve for example, problems in MRI reconstruction so you all familiar with MRI right there is a source which takes a cross-section of your body image they use these kinds of properties to derive the images of your body okay there's another final property about these two which is also very nice can someone tell me what that is will be Correct correct that is right. So, if I take all the if I take the row space and if I take the null space together, they span the entire $\mathbb{R}^n$ right.

So, this is equal to $\mathbb{R}^n$ and this whole thing is equal to $\mathbb{R}^n$ ok. So, I can take a vector any vector writes it as having a row space component and a null space component ok. Any questions on this so far? So, the next moving on from the fundamental subspaces of a matrix, the next thing comes to the Eigen and SVD decomposition of a matrix ok. So, let us start by looking at eigenvalues ok. So, Eigen decomposition.



I mean this is really obvious I really do not need to define it, but I will still say it. We all know what is the eigenvalue problem?

$$Ax = \lambda x$$

where λ is a scalar, actually need not be it can even be complex valued. You can have a real-valued matrix with a complex eigenvalue it can happen and this is obviously only valid for what type of matrices? square matrix right. So, A is square. Otherwise, this will become an illegal operation right $A \cdot x$ will give you the wrong dimension.

So, A is square ok. So, in optimization there is a special class of matrices that comes again and again which is a symmetric matrix ok. So, A let us start with A symmetric, obviously if it is symmetric I am implicitly assuming that this is a square matrix ok. If A is symmetric there is a very nice result which states that a property of the eigenvalues for a symmetric matrix the eigenvalues are real valued eigenvalues right the eigenvalues are really ok and let us say that these are λ_i's ok. Corresponding to each λ_i there will be some eigenvector right eigenvectors ok q_i.

Now, for a when A is symmetric you get the very very famous eigen decomposition of a matrix right. So, the eigen decomposition in the case of a symmetric matrix is going to be written as I am going to write it in a form that you might not be familiar with first. So, what does this look like? I am writing a matrix A as a sum of something, but let us look at each one of these quantities first. Sorry yeah, I got this in the wrong place yeah. What does this look like? This is not the inner product, the inner product between two vectors was $x^\top y$, but this is more like $xy^\top$.

So, this is called the outer product. which is a very powerful way of writing a matrix and summed over all of the eigenvectors and eigenvalues right. You can see that $q_i q_i^\top$, what size will it be? q_i is of what size? $n \times 1$ $q_i^\top$ is $1 \times n$. So, $q_i q_i^\top$ is $n \times n$ right. what is again those of you who are familiar with linear algebra, what is the rank of $q_i q_i^\top$? It is rank 1 ok.

You will see all the columns or all the rows are just linear multiples of each other right. So, it has just rank 1. So, in this notation you see that I can write a matrix as a sum of rank 1 matrices. This is again something used extensively in optimization, machine learning all of these things particularly where there is the application of sparsity happening, but we will hopefully talk about that later in the course. So, this is one way of writing it there is the more common way of writing it is to write it in the form of product of matrices which is going to be

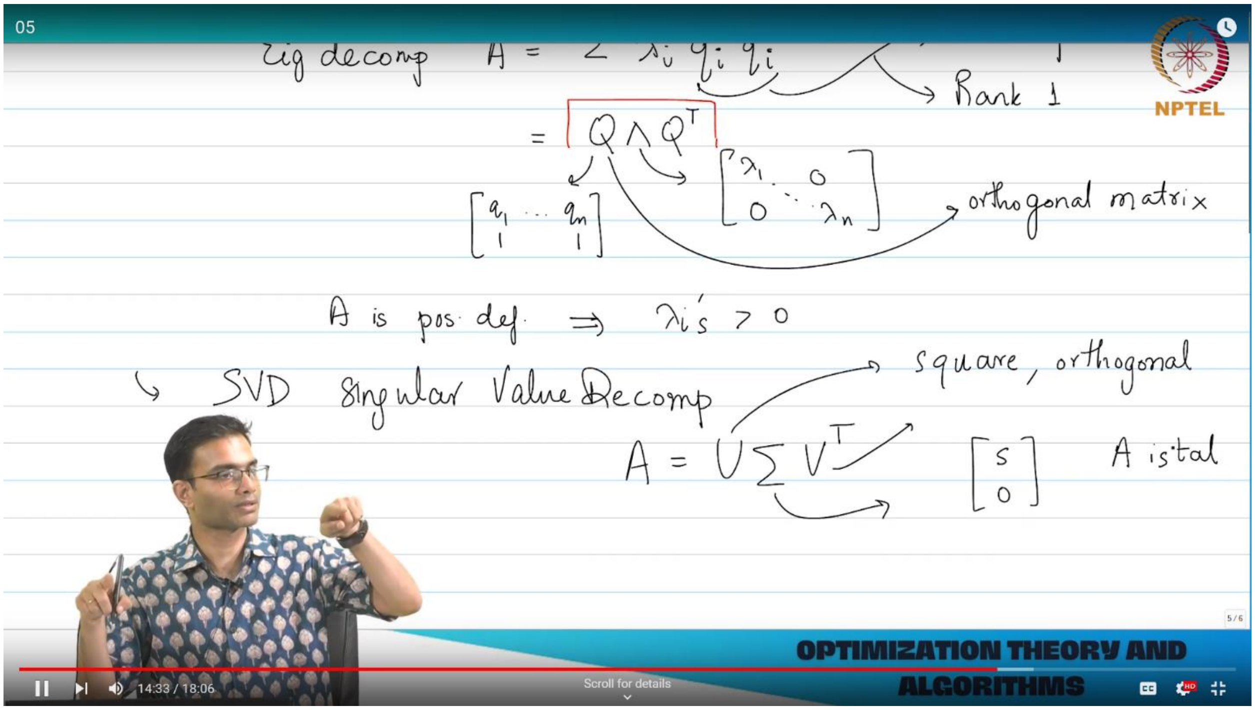
$$Q\Lambda Q^\top$$

ok. What is Λ? Capital Λ, it is a diagonal matrix with ok and Q is simply q_1 up to q_n vectors like this. You can verify that once you open this up you get the same expression over here ok. Is there something, there is something further special about this Q? Can someone remember what it is? It is an orthogonal matrix, which just means that each column is orthogonal to each other. One further simplification happens if matrix A is positive definite, what happens to the eigenvalues? They become positive ok. So, this was about your eigenvalue decomposition.

Particularly, I mean eigenvalue decomposition is used not just extensively in electrical engineering, but those of you who are interested in quantum mechanics will find these matrices appearing everywhere. It is just the nature of physics that you end up with symmetric matrices. So, this was the EVD or eigenvalue decomposition. The other decomposition which is of you know bread and butter for electrical engineers is actually not just electrical, all of engineering is SVD, singular value decomposition. And it simply writes that notice the similarity with the eigenvalue decomposition.

Eigenvalue decomposition was very nice square matrices appearing everywhere, but in real life your matrix A may not be square right. So, you are going to write this as $U\Sigma V^\top$ and something special about U and V they are square they are orthogonal. Square and orthogonal. People will refer to U and V as left

singular vectors and right singular vectors and what about this capital Σ? It has singular values on the diagonal it need not be square. In fact, if A is rectangular Σ is going to be rectangular right.



So, what are the choices like? If A is for example, a tall matrix So, you will have Σ will be like this, right? So, S is a diagonal matrix $\sigma_1\sigma_2$ up to σ_n right. So, I will write it out fully. This is if A is tall. S if A is square, and S and 0 if A is fat ok and S is simply a diagonal matrix of singular values ok. Now, the convention that we will follow in this course which is a very common convention in engineering is σ_1 is the highest singular value and σ_n is the lowest singular value ok.

So, if I do not say anything assume σ_1 is the largest σ_n is the smallest. You notice that I can rearrange these σ's and get a new decomposition, but it really does not matter right. So, this is always going to be largest, for example, if you use MATLAB, it will give you the decomposition with σ_1 being the largest. And this singular sorry the matrix condition number which I had defined a few minutes ago, $\kappa(A)$, you can see that it you get a very nice expression it ends up being $\frac{\sigma_1}{\sigma_n}$, ratio of the largest to smallest singular value ok. So, this is something which again in MATLAB is just a simple command called 'cond', you give it like this and it will tell you what is the condition number ok.

What is the condition number of the identity matrix? 1 right. So, condition number actually has a property that the best you could do is a condition number of 1 ok and the worst is ah well it is unbounded ok. Additional review I am pretty sure all of you know this that ah for a square matrix When I say $\text{trace}(A)$ does everyone know what that means? Sum of diagonal elements. Is there any relation of that with the eigenvalues? It is the sum of eigenvalues right. So, this is $\sum A_{ii}$ over i which is also the same as $\sum \lambda_i$ ok.

Any other relation between some matrix property and eigenvalues? Determinant right. So, the determinant of A is the product of the eigenvalues. So, that kind of brings us to an end of the review of linear algebra ok. So, this is for you to kind of go back and look how familiar are you with it is it something that you can pick up by yourself or you need to actually do the whole course yeah that is something for you to assess.