Applied Linear Algebra
Prof. Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras

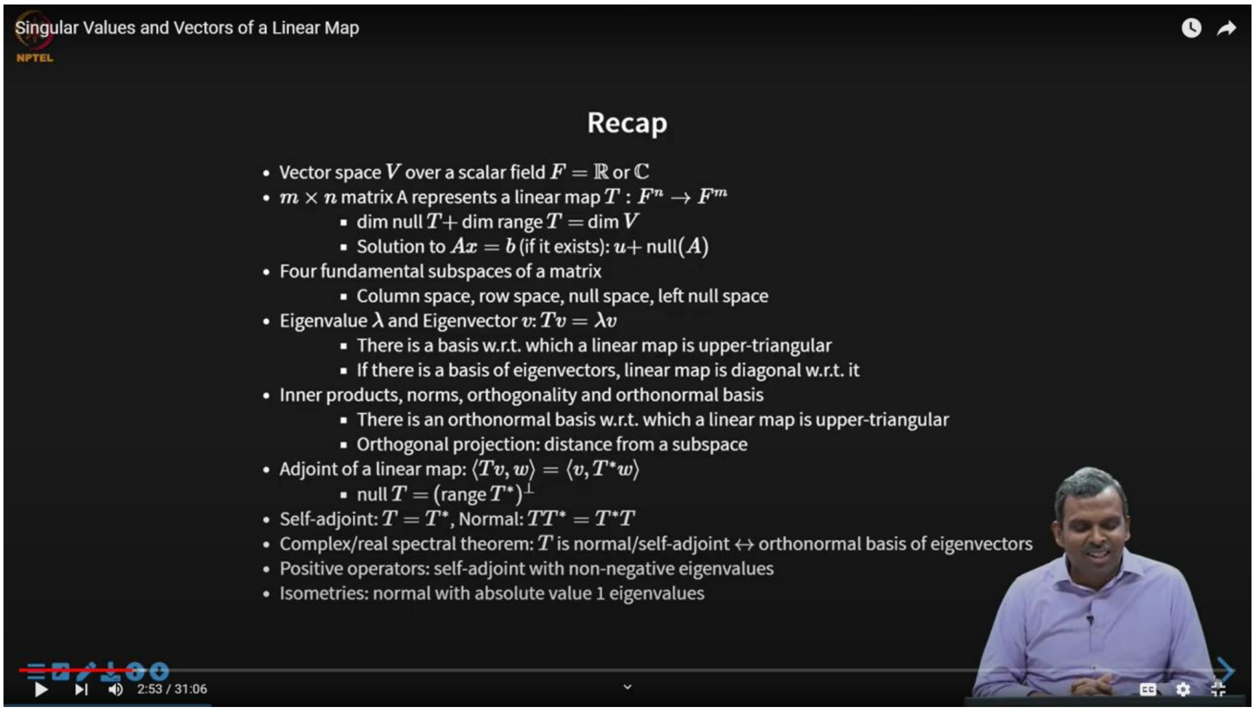
Week 12
Singular Values and Vectors of a Linear Map

Hello and welcome to this lecture. This is the last week of this class recording and we've seen quite a bit of interesting ideas on how to study vector spaces, how to study linear maps, right? So linear maps form a huge part of the entire area. And the more we study, the better off it is, right? Better off we are in our understanding. So we saw for instance eigenvalues played a very important role in the study of linear maps, right? So they gave you one dimensional subspaces which are invariant and that gave you a pretty good idea of what a linear map is. In particular, if you had enough of these one dimensional invariant subspaces, if your operator became diagonalizable, then, you know, there was a very simple diagonal form in an eigenvector basis and we saw, you know, special type of operators like self-adjoint, normal operators for whom there's an orthonormal basis of eigenvectors. So those operators are very special because they become diagonal over an orthonormal basis. Now there is an equally powerful idea with huge applications today which is not based on eigenvalues, it's based on something called singular values and singular vectors. So like you had eigenvalues and eigenvectors, it is possible to define a notion called singular values and singular vectors for in fact the general linear maps, right?

So we saw values eigenvectors are for operators, right? And now one can define these singular values and singular vectors and these singular vectors always end up being orthonormal, that's very nice, and singular values always end up being non-negative and they are there for anything. For linear maps, any map, whether it's from $V \to W$ or operator $V \to V$, you can define all that. So it's a different way of looking at a linear map, right? So eigenvalues and eigenvectors tell you, you know, one dimensional invariant subspaces. That's one way of analyzing it. This singular value, singular vectors are another way of looking at it. You will see it's very interesting and very surprising how simple and elegant these singular values and singular vectors are and we will do that. In the first lecture, I will simply introduce what is a singular value, what is a singular vector. We won't prove some important results, we will do that in the subsequent lecture. But it is important to get a feel for what it is and what is the difference between this and eigenvalues, eigenvectors and see some comparisons. So we'll do that in this lecture. So let's get started.

There's a quick recap of all that we've done so far. It's nice when the entire course is recapped into one page, you feel like you've got a good sense of what's going on. So we've been looking at vector spaces over real or complex field. We saw how a linear map has a matrix representation, the fundamental theorem of linear maps which is, you know, dimension of null plus range equals dimension of V, it's very important. We solved linear equations using the null space idea, found all solutions for it and these four fundamental subspaces of a matrix give you an idea of how to think of the linear map and all that. And then the important idea of eigenvalue and eigenvector. So in particular if the matrix can become diagonal with respect to an eigenvector basis, then things are really simple particularly with respect to powers of the matrix, right? Powers of the operator. So it is very useful to have the eigenvalue eigenvector notion.

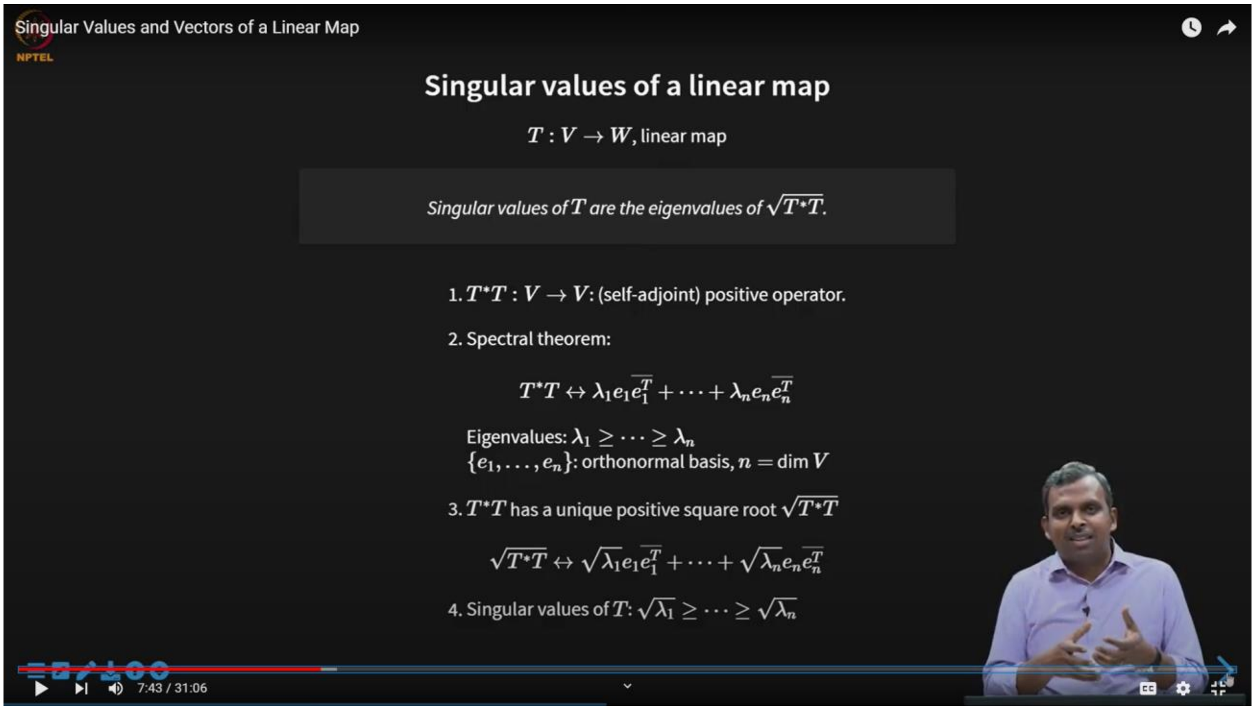
(Refer Slide Time: 02:53)



And then we saw inner products and orthogonality, orthonormal basis. This orthonormality simplified a lot of things for us including basis representation and all that. And then the important idea of adjoint and how this adjoint sort of preserves, tells you how inner products work along with linear maps in a very clean way. In particular that led us to these wonderful operators called self-adjoint and normal operators which have orthonormal eigenvector bases, okay? So that is very powerful and strong. And then we saw special cases, more special cases called positive operators. And in fact even more special cases called isometries. These are special type of normal self-adjoint operators and they have their connections to, you know, quadratic forms and all that. Particularly the positive operators. Isometries we'll see some connections today. And then, not, maybe not today... In these singular values and singular vectors, isometries will play some role, we will see some interesting ideas from that, okay? So this is a quick recap. Let us jump into singular values, okay?

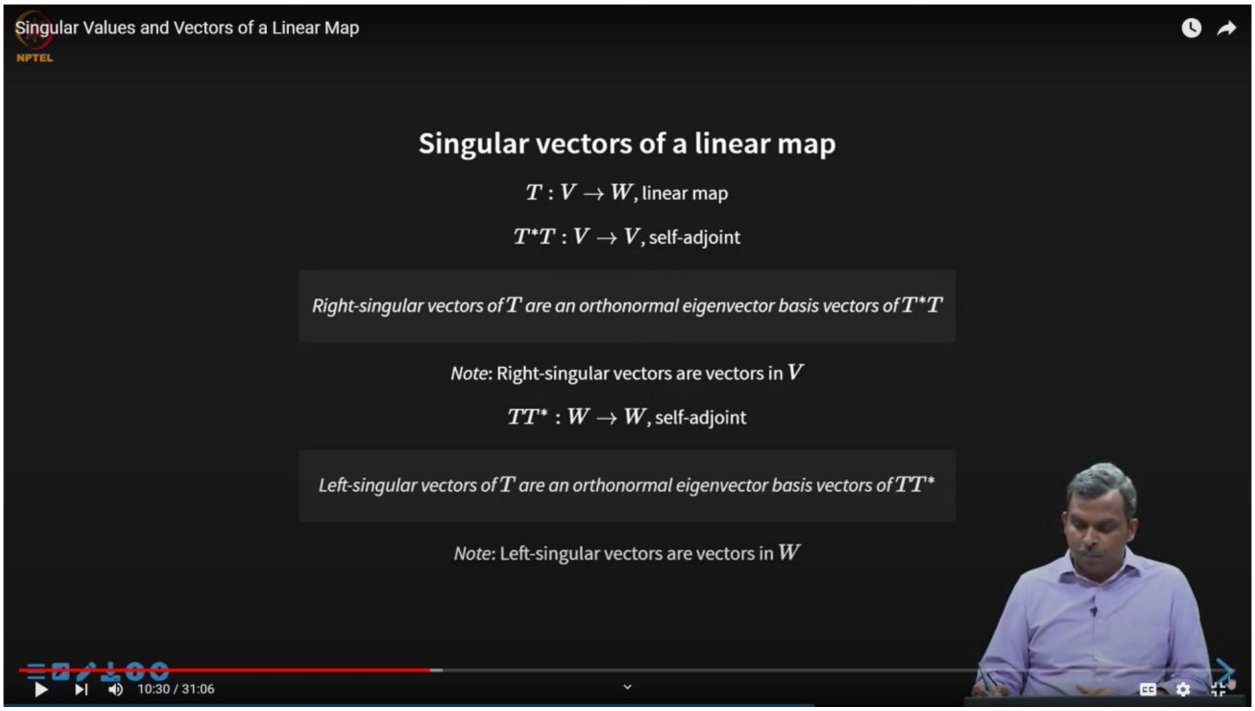
So here's the definition of a singular value of a linear map, okay? So let's take a linear map $T: V \to W$, vector space V to vector space W. The singular values of T are simply the eigenvalues of $\sqrt{T^*T}$, okay? So that is a pretty big operator written in a very compact way. $\sqrt{T^*T}$, okay? So T^* is the adjoint of T and this is the square root. So this is, I mean quite a few interesting properties in just this definition. It's an easy enough definition. But quite a few interesting properties. First property is - this a properly well defined thing. I mean it's not something wrong. $T^*T: V \to V$ is a self-adjoint positive operator. Positive is only for self-adjoint so we know that's okay. So square root and all exist. But before that let's just use the spectral theorem. When you use the spectral theorem, you're going to be able to write T^*T over this suitable basis in this very simple form, right? $(\lambda_1 e_1 e_1^* + \cdots + \lambda_n e_n e_n^*)$, okay? And these eigenvalues are non negative, so you can sequence them in, you know, decreasing order like that. Then this e_1 to e_n is an orthonormal basis. n is the dimension. So we know this is true. From here we can define the unique positive square root, $\sqrt{T^*T}$. I never in the lectures proved the uniqueness of this. There is a proof in your book, you can go take a look. It's a unique square root. So we'll pick up the unique positive square root, $\sqrt{T^*T}$ which would have, you know, the basis representation in, matrix representation written in this form. So wherever you have lambda, you put square root, okay? So everything else remains the same. So the singular values of T in this, written down in this way are simply the $\sqrt{\lambda_1}$ and that would fall in decreasing order, go all the way to $\sqrt{\lambda_n}$, okay? So this is a picture of how singular values are defined. We will see soon enough a couple of examples in this lecture itself, it will be clear to you how this works. But this is the definition and this definition makes sense, okay?

(Refer Slide Time: 07:44)



So why is it that singular values describe T? So it looks like singular values are numbers related to some other maybe operator $\sqrt{T^*T}$, right? So why should that be connected to T? It turns out there is a very strong connection. We will prove it eventually in the next lecture or so. But in this lecture I will simply illustrate that connection with some examples, okay? But anyway, I mean you can also imagine why this T and T^* have been, I've been hinting at how T and T^* are sort of the same. So when you take $\sqrt{T^*T}$, you sort of expect, you know, something similar to T in some sense, right? So a very rough high level picture. But that's sort of the reason why this T and $\sqrt{T^*T}$, they have sort of an intimate connection which we will establish later on. But for now it's not surprising that these two are related, okay? That's that for now. We will pick it up later.

(Refer Slide Time: 10:30)

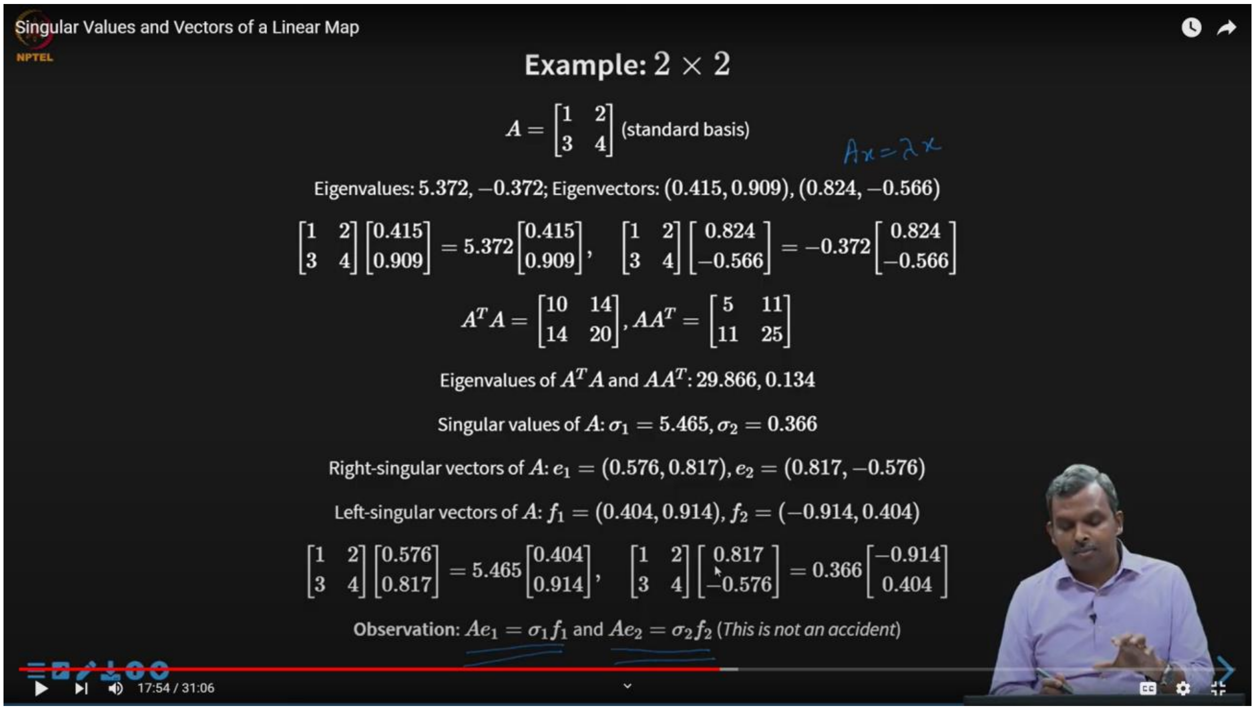


Okay. So what about singular vectors of a linear map? These are also very easy to define. Once you know the singular values, and when you see this T^*T connection, you will see that it is easy to define. You once again have a linear map $T: V \to W$, you can define T^*T which will be from V to V. It is a self-adjoint operator. The right singular vectors of T what I call as right singular vectors, these are simply an orthonormal eigenvector basis for T^*T, okay? So you pick whatever orthonormal eigenvector basis you want for T^*T, those would be called as right singular vectors. So it's sort of an easy definition in some sense, you know? Singular values are eigenvalues, singular vectors are simply the corresponding eigenvectors, right? The $\sqrt{T^*T}$ are, okay, right? T^*T and $\sqrt{T^*T}$ have the same sort of eigenvector basis in some sense, right? So it's okay to define it with T^*T in this case, okay? So remember right singular vectors are vectors in V, okay? So that is something important to remember. So now you can see where the left singular vectors will come from. Instead of T^*T, you look at TT^*, okay? Remember T^*T is self-adjoint, the adjoint of T^*T is T^*T itself. So TT^* is a different other thing, right? So if T were an operator and T were normal, TT^* and T^*T will be the same. But they need not be the same, okay? TT^* and T^*T can be different. But of course, if it is a linear map, general linear map with V and W being different, then TT^* and T^*T can have no chance of being the same, right? So because T^*T is from V to V, TT^* is from W to W. But once again it is a self-adjoint operator, so I can meaningfully think of an orthonormal basis of eigenvectors. So if I take an orthonormal basis of eigenvectors of TT^*, I end up getting what are called left singular vectors of T, okay? So this is the definition. Once again, right singular vectors are vectors in V, left singular vectors are vectors in W, they are always orthonormal, okay? By definition they are orthonormal, there is no, they cannot be anything else. And they are very easy to come up with, okay? So even for a linear map T when V and W are different, these singular vectors are very well defined. They end up being orthonormal basis vectors of V and W related to T^*T and TT^*, okay? So simple definition for singular values and singular vectors.

Let's see a couple of illustrations and computations and see some surprising properties which we'll eventually prove later on, okay? So the simplest example is this 2×2 and I've been doing this $[1\ 2; 3\ 4]$ throughout this course, so we can pick up the $[1\ 2; 3\ 4]$ again, okay? So you have a matrix A, 2×2 which is $[1\ 2; 3\ 4]$. Let's say standard basis, okay? That's an easy thing to start, okay? We know that this matrix itself has eigenvalues and eigenvectors, you can compute the eigenvalues and eigenvectors. I have done them, I have done that here. So you see that you get two distinct eigenvalues $5.372, -0.372$ that's what it seems like. I've calculated it. I've also calculated the eigenvectors, that's, one eigenvector is $[0.415; 0.909]$, the other eigenvector is $[0.824; -0.566]$. So these are linearly independent, these two eigenvectors. But they are not orthogonal. You can multiply and check that it won't work out to be orthogonal. So non-orthogonal to be expected, right? A is not a symmetric matrix, so you will have a non-orthogonal set of eigenvectors. So what is the property of eigenvectors? You can check this property. So the matrix itself A multiplied by the eigenvector is going to give you eigenvalue times the same vector. You can check that this is true, right? So it's easy enough to see. So this is the structure of the eigenvalue eigenvector.

Now I want to contrast this with the singular value singular vector picture, okay? So this picture itself is quite useful. We've seen that it helps you. So in fact you can write A as $S^{-1}DS$, okay? You can write with these exact eigenvectors, you know how to form the S matrix. And A will become $S^{-1}D\,S$, right? So D becomes diagonal. And so it's useful, this eigenvector eigenvalue representation is very useful even for this matrix. So now let's see how the singular values singular vectors look, okay? So we know we have to do A^TA and AA^T to go to the singular values singular vectors. If you do A^TA, you will get a symmetric matrix. $[10\ 14;\ 14\ 20]$. You can verify that. AA^T ends up being $[5\ 11;\ 11\ 25]$. I hope this is correct, not made a mistake here. Hopefully AA^T, yeah, looks okay. Okay? All right. So I was just checking to make sure, yeah, I think it's okay. So it's $[5\ 11;\ 11\ 25]$, okay? Even if there is some mistake, it's okay, illustration is more more important, okay? So the eigenvalues of $A^T A$ and AA^T, okay? We know that this is AB, BA, right? So the eigenvalues, we know all the non-zero eigenvalues have to be exactly the same, there can be more zero eigenvalues depending on the dimensions of $A^T A$ and AA^T. In this case both of them are same dimension 2×2, so the eigenvalues will end up being exactly the same. So you do $A^T A$, AA^T, you will get the same eigenvalues 29.866 and 0.134. So this is a trick you should keep in mind. So sometimes, you know, if A and A^T are not of the same size, $A^T A$ might be much much smaller than AA^T. So when you want to find eigenvalues, you should be smart enough to go pick the smaller one and try to find it, okay? So that is something useful. But this is easy enough to see, okay? So once you have the eigenvalues of, say, $A^T A$, I can find the singular values. Singular values are simply the square root of the eigenvalues. So you take $\sqrt{29.866}$, you get 5.465. $\sqrt{0.134}$, you get 0.366. So that's the singular value.

(Refer Slide Time: 17:57)



It's typical to denote singular value with this σ, okay? σ_1, σ_2 in that sequence. $\sigma_1 \geq \sigma_2$, etcetera, okay? One can also find right singular eigenvectors and left singular eigenvectors, okay? For these two eigenvalues. In this case it's all unique and all that so you'll get the same answer. You may change up to sign, but only up to sign you can change, everything else will work out and it's very typical to write right single eigenvectors, left single eigenvectors as orthonormal basis, okay? So you can see e_1 and e_2 is an orthonormal basis for, you know, 2 dimensional $\mathbb{R}^2$ and similarly f_1 and f_2 is also orthonormal basis, okay? e_1 and e_2 are eigenvectors of $A^T A$, f_1 and f_2 are eigenvectors of AA^T, okay? So these are calculations you can check. I am not going to go through great detail for you. You can check that these are true, okay? So you see eigenvalues were defined in one way. You know how to calculate them, singular values and singular vectors are ending up in another way. I know that eigenvalues gave me the, you know, $S^{-1}DS$ type transformation for A, representation for A, and that helped me in so many ways. What is going to happen here, okay?

So we will see what it does eventually, but let us look at something very interesting. So when the matrix operates on the eigenvector, you expect something very simple in terms of the eigenvalue and itself, right? So what happens when the same matrix A operates on the right singular vector, right? So let us pick the right singular vectors and say the matrix A operates on on them. Look at what happens. Some magic will happen here. So you take A and multiply by the right singular vector, the first one here e_1. e_1 ends up, notice what you're getting, you will get σ_1 the singular value multiplying f_1, okay? And if you take A and multiply, I mean multiply on the right with e_2, right? Ae_2, the second single right singular vector, you end up getting the second singular value multiplying the second left singular vector, okay? So these kind of properties, this is not an accident. $Ae_1 = \sigma_1 f_1$ and $Ae_2 = \sigma_2 f_2$. This is the defining property for singular vectors and singular values, just like you have, you know, $Ax = \lambda x$.

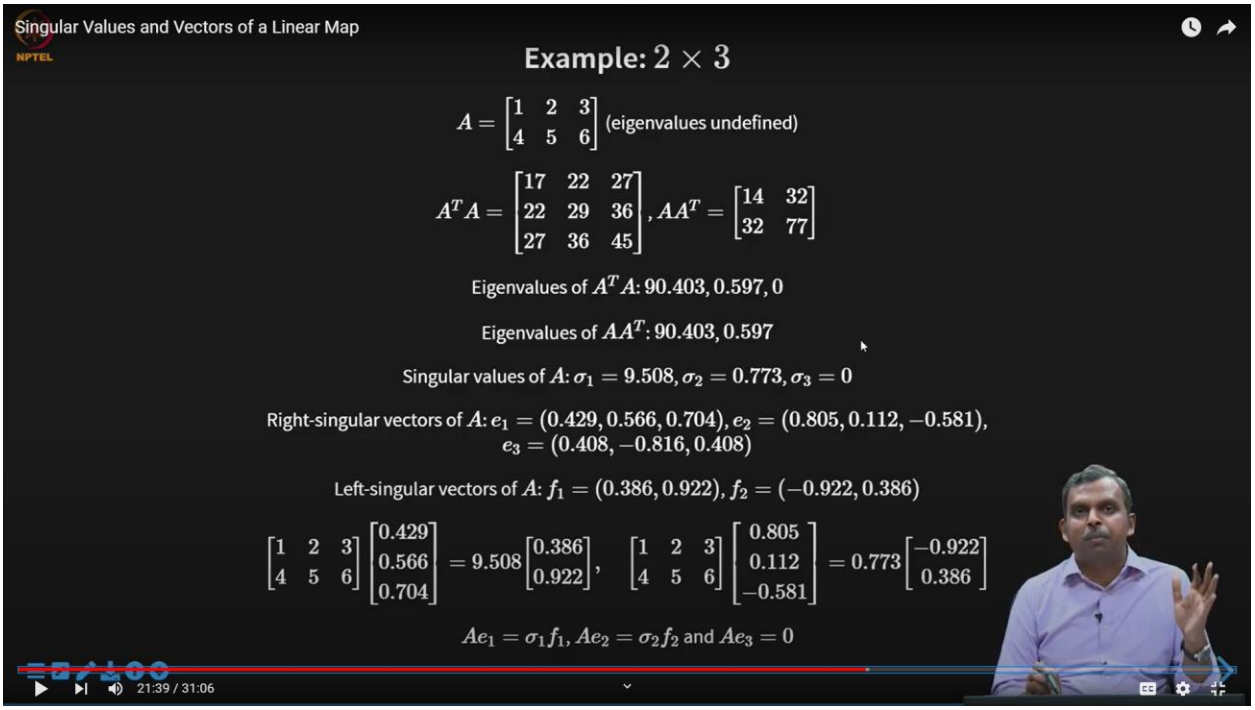
Maybe I should write that alone. So $Ax = \lambda x$ is sort of the defining property for eigenvalues definition. This is the defining property for a singular value singular vector definition, okay? So you have, you have always orthogonal. So here you have two eigenvectors, they are not orthogonal. Here the singular vectors are always orthogonal. And then what happens to singular vectors when A operates on them? They go to orthogonal vectors as a result, right? The output is still an orthogonal set of vectors. These two are orthogonal, they are scaled by the singular values and that's it, right? So it's not like it's invariant. The singular vectors are not invariant, but they sort of provide a very nice and complete picture of what A is, okay? They take an orthonormal basis to an orthonormal basis in some sense, right? So that's the crucial idea. We will prove this eventually, in the next lecture or so we will prove this result and this is the essential aspect of right and left singular vectors and singular values, okay? So it's important to understand.

This is one example, okay? Hopefully this was clear. Let's see one more example just to drive home the point of how singular values and singular vectors work and this time I will take a non-square example 2×3, okay? So if you take 2×3, then there is no question of eigenvalues or anything, right? So those things do not make sense. So we have to go to singular values. When we go to singular values, let's look at A^TA and AA^T, okay? So here you see the difference, right? A^TA is 3×3, AA^T just 2×2. And you can look at eigenvalues of A^TA, you will get some number, 90 and some 0.5. Eigenvalues of AA^T we know what it will be, right? It will be, the two non-zero ones will be the same, the zero will drop out, right? That's what will happen. So the singular values of A are in fact the eigenvalues of A^TA. So it is common to, so you can take, I mean square root of the eigenvalues of A^TA. So you get 9.5, 0.773 and then the σ_3 is zero, okay? So you get that and you can find right singular vectors of A. This will be the eigenvectors of A^TA. You will get e_1, e_2, e_3. And once again e_1, e_2, e_3 is an orthonormal basis for $\mathbb{R}^3$, okay? So that's $\mathbb{R}^3$, that is, you know,

V to W, right? A is from V to W. V is $\mathbb{R}^3$. W is $\mathbb{R}^2$. So e_1, e_2, e_3 is an orthonormal basis for $\mathbb{R}^3$, okay? So this corresponds to the eigenvectors of $A^T A$, the orthonormal basis eigenvectors of $A^T A$.

Similarly, you can also find left singular values of A. Now this will be now left singular vectors of A. I'm sorry, these will now be an orthonormal basis for $\mathbb{R}^2$, okay? So this is different, right? The range is in a different vector space. But still the orthonormality will be satisfied, okay? So this is sort of gives you an idea of how it works. So even when eigenvalues are not properly defined, you can say something with singular values, okay? So again you test this little result that we have. Ae_1 is again $\sigma_1 f_1$. Ae_2 is $\sigma_2 f_2$. Ae_3 is 0, okay? So you notice what's going on here. So there is an orthonormal basis which transforms through A to another orthonormal basis and the singular values are the scaling factors that come in the middle, okay? So this is a very powerful description for any linear map. So is it true for any matrix, okay?

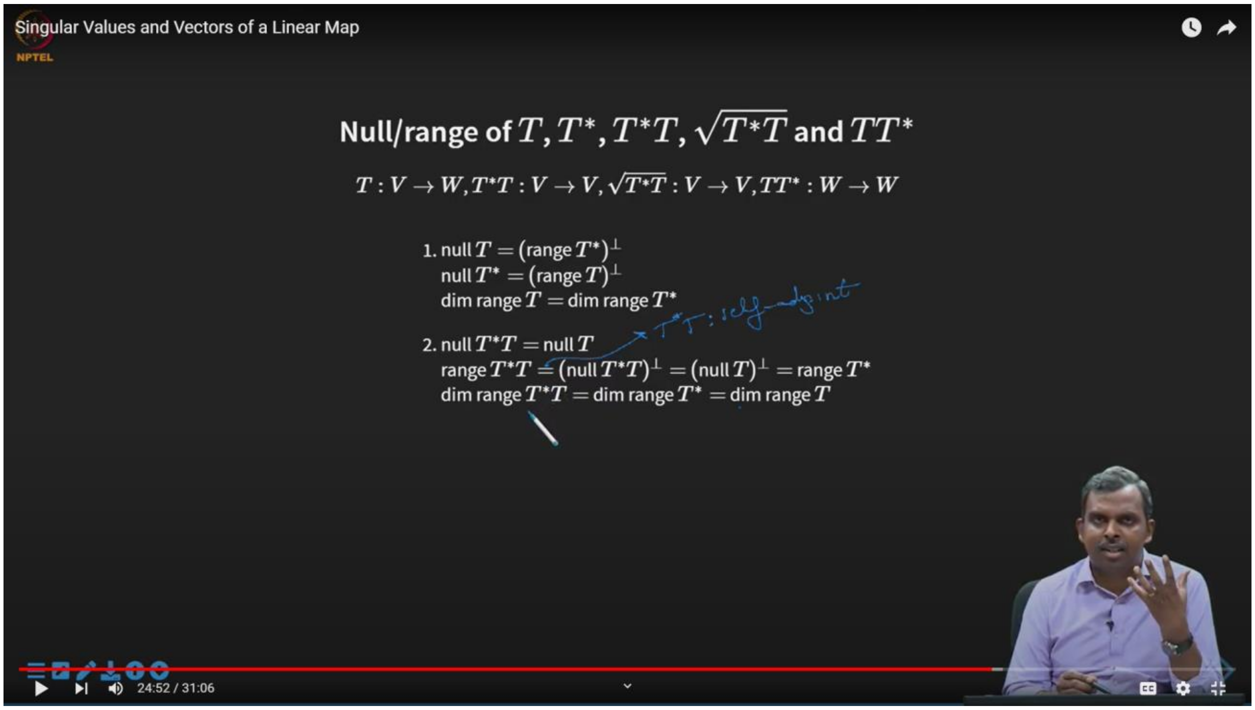
(Refer Slide Time: 21:39)



So these are the questions we have to ask. We'll ask this and we'll answer these questions in the next lecture. But in this lecture this, I mean I hope you see the magic of what's going on. So we are able to find an orthonormal basis which when hit with a matrix A takes you to an orthonormal basis. That ends up being the huge simplification here and it's all connected to this $A^T A$ and AA^T and its eigenvalues and eigenvectors, okay? So hopefully these two examples give you a feel for how to work with singular values and singular vectors. And always remember this $A^T A$, AA^T, one of these will be smaller, okay? So you should just go there, go to that and then go to the smaller one and then work with that, okay? So work with that eigenvalue and those eigenvectors and then

f_1 and e_1 you can find through these relationships, right? So it's all easy to sort of do. You don't have to do a lot of work for these small size problems. But of course today SVD is one of the, is the heart and soul of any numerical implementation of linear algebra. In fact all problems, any linear algebra thing that you ask a computer program to do with a matrix, it will first find SVD, okay? And all answers about the linear map can be given from the SVD, okay? So you'll see why. I mean, we will see later on in more detail why it's, this SVD is so powerful. Singular value decomposition and singular values and singular vectors are so powerful. And they give you a very succinctly nice description of any linear map, okay? So this is just a numerical illustration at this point. We will see more proofs and more properties later on, okay?

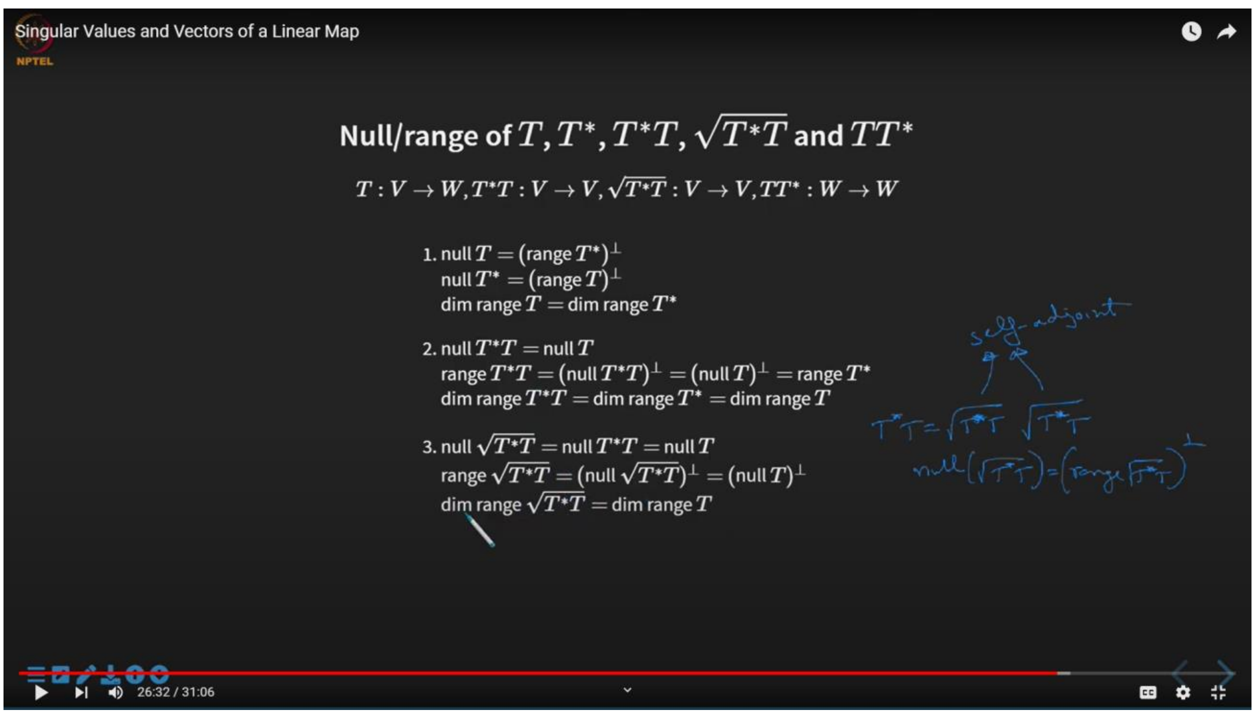
(Refer Slide Time: 24:54)



So at the heart of all of this is this. All sorts of connections between null space, range space of T, T^*, T^*T, $\sqrt{T^*T}$ and TT^* , okay? So these are all the various things that are going on here and they are from various spaces to other, that's written down here. And the nulls and ranges are all strongly coupled and connected, right? And we've seen how to work with these before. The whole bunch of results we saw on what happens when you multiply two operators, right? When you compose two operators, how do you think of the null space of that and all that. But, you know, T and T^* have a very special connection. So you can use that and then simplify all the properties. So I am going to write down four different properties. They may be a little bit confusing, but think about them clearly, you will get a good idea and that is very critical in our proofs and other things that we do, okay? So this is important to understand.

So the first is T and T^*. This is standard. right? So any T and T^*, null T is $(\text{range } T^*)^\perp$, okay? And the same thing with T^*. Instead of T if you put T^*, you will get null T^* is $(\text{range } T)^\perp$. And from this you can conclude that range of T and range of T^* have the same dimension, okay? This is true for any operator T, we know that this is true, right? So once you have something like this, it turns out... See null of T^*, null of T is range of T^*. So if you look at T^*T, right, the only way to get the null of thing is to make T itself null because, why? Because range of T and null of T^* have a trivial intersection, right? So range of T and null of T^* are actually orthogonal complements, so they do not intersect at all. So the null of T^*T is going to be equal to null of T. So this is a crucial relationship. So because of this, range of T^*T which will have to be null T^*T orthogonal complement, right? This, why is this true by the way? The first part, why is this true? This is true because T^*T is self-adjoint, okay? So T^*T is self-adjoint. Once you have a self-adjoint operator, its range is the orthogonal complement of its null. Now what is null T^*T? It is the same as null T. So $(\text{null } T)^\perp$, what is $(\text{null } T)^\perp$? From here, range T^*, okay? So range T^*T and range T^* are exactly the same, okay? All right? They are equal. There's not, there's no difference here, okay? It's not just dimension or anything, they are actually equal. Now if you look at dimension, so dimension of range T^*T becomes dimension of range T^* and that is equal to the dimension of range T also, okay? So in dimension, all of these things are the same. So that's, it adds to my notion that, you know, T, T^* and T^*T, TT^*, share a lot of these important properties. And that's why it's interesting to look at them, okay?

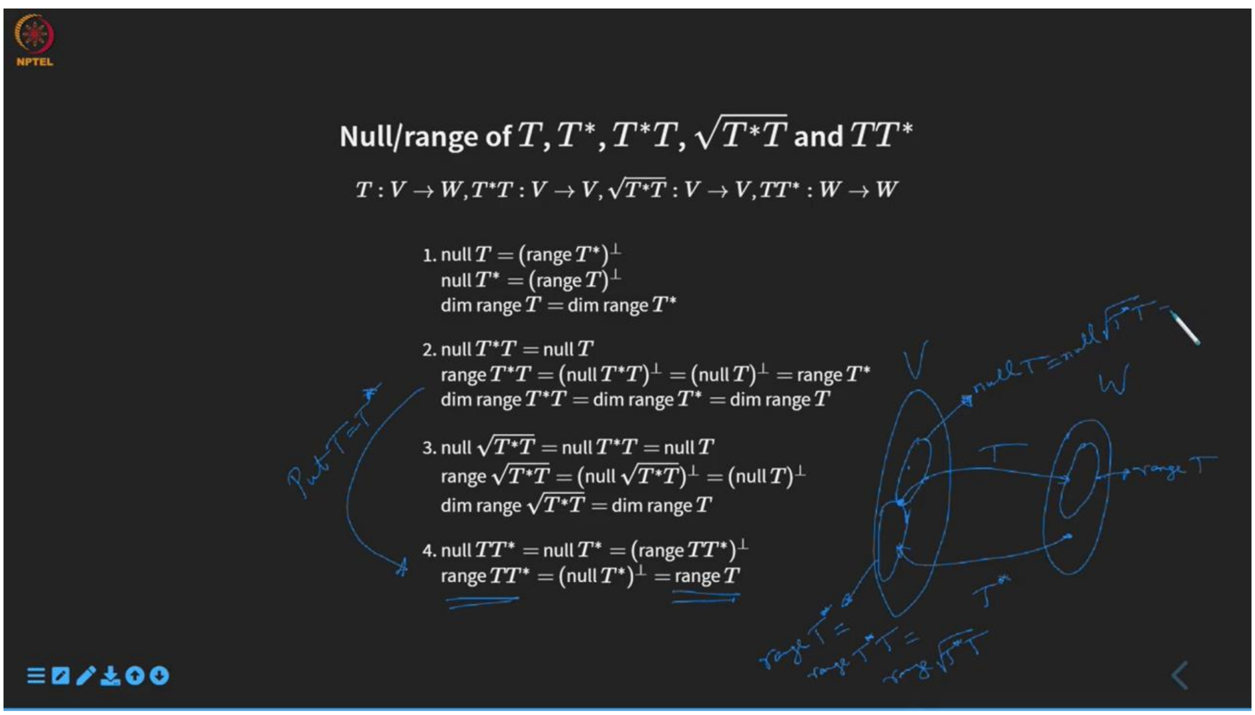
(Refer Slide Time: 26:32)



All right. So this is the second one. Let's go to the third one which talks about $\sqrt{T^*T}$. So now if you look at $\sqrt{T^*T}$, how do you sort of see this? So you see that, you know, T^*T equals $\sqrt{T^*T}$ times $\sqrt{T^*T}$. Now all these are self adjoint, right? So these two are both self adjoint. So what will happen? The range of T^*T and null of $\sqrt{T^*T}$ will be orthogonal complements, right? So at this point if you see, null of $\sqrt{T^*T}$ is the $\left(\text{range } \sqrt{T^*T}\right)^{\perp}$, okay? So because of this, right? The only way to make T^*T null is to also have $\sqrt{T^*T}$ null, going to the null space, right? So because the range of this and the null of the next one will have no intersection, have a trivial intersection, okay? So that is why null of $\sqrt{T^*T}$, because of the self-adjoint property is equal to null of T^*T and null of T^*T is the same as null of T, okay?

So range is also the same, right? So dimension… Means same in the sense, you know, very similar description to before. So dimension of range of $\sqrt{T^*T}$ is also equal to dimension of range of T, okay? So this is important to know, okay? So T^*T, $\sqrt{T^*T}$, all of them have the same range dimensions, okay? And this will be useful for us later on, okay? So hopefully that gave you a clear idea.

(Refer Slide Time: 29:52)



So finally what about TT^*? So far I've been talking about T^*T. If you want to do TT^*, simply take any of these relationships and instead of T you replace with T^*, right? So that's the idea. You do that. So how do you get from here to here for instance? Put T equals T^*, okay? So you replace T with T^*, you will get null of TT^* is null of T^* and that's (range TT^*)$^\perp$, right? So null and range are the same. So range TT^* is $(\text{null } T^*)^\perp$ and that's range T, okay? So this is also important. Range TT^* is range T, okay? So just like we had here, right? Range of TT^* is the same as range of T, okay?

So if you want more pictures… So let us say this is V and this is W. T takes you from here to here. T^* takes you from here to here. So range of T will be somewhere here, right? So this is range of T let us say. What is T^*, and then T? So if you take T^* and then T, okay, so you will have a range of T^* here. So these two have the same dimensions. Maybe I didn't draw it correctly. Maybe I should do that. Just cut it short a little bit so they have the same dimension. These two have the same dimension. So if you do this, T^* and then T, you go back to the same range, okay? So the range of T^*T is the same as range of T, okay? And then look at $\sqrt{T^*T}$, okay? So if you look at range of $\sqrt{T^*T}$… So you have null T, right, null T, okay? So null T, okay, null T and range T^* are orthogonal complements. So they allow only a trivial intersection, right? null T. And null T is the same as null of $\sqrt{T^*T}$, okay? So the $\sqrt{T^*T}$ is both of these are same.

And you know, right, so null is the same. So if you want to look at the range, you will end up getting the range of T^*, right? So if you take the orthogonal complement of this… So range of T^* is the same as range of $\sqrt{T^*T}$, right? Range of T^* is the same as range of T^*T is the same as range of $\sqrt{T^*T}$, okay? All of these have the same range. You can also write similar expressions for range of T and range of TT^*, okay? So they all have the same range, they all have the same null, okay? So this is also equal to null of T^*T, okay? So hopefully you see the picture. These are two orthogonal complements and they all are the same, okay? So as far as sizes and everything is concerned, they are the same. The dimensions are the same, okay? And they also have more intricate connection, particularly the square root. I will show a very very close and intricate connection in the next slide and that will give you what is called the singular value decomposition and that is very very powerful and finally it will connect all that we've been seeing as illustrations in this class, okay?

So this slide was maybe a little bit confusing. But it’s sort of important to know how all these null spaces and range spaces are very directly and closely connected in terms of just being equal and dimensions being same and all that, okay? So I’ll stop here. We saw two or three nice illustrations of how to compute singular values, how to compute left singular vectors, right singular vectors and the wonderful relationship between them, okay? A times a right singular vector is singular value times a left singular vector, okay? So that's a very critical relationship. And singular vectors are always orthonormal and that will lead us to what's called singular value decomposition and polar decomposition. And we'll see that in the next lecture. Thank you very much.