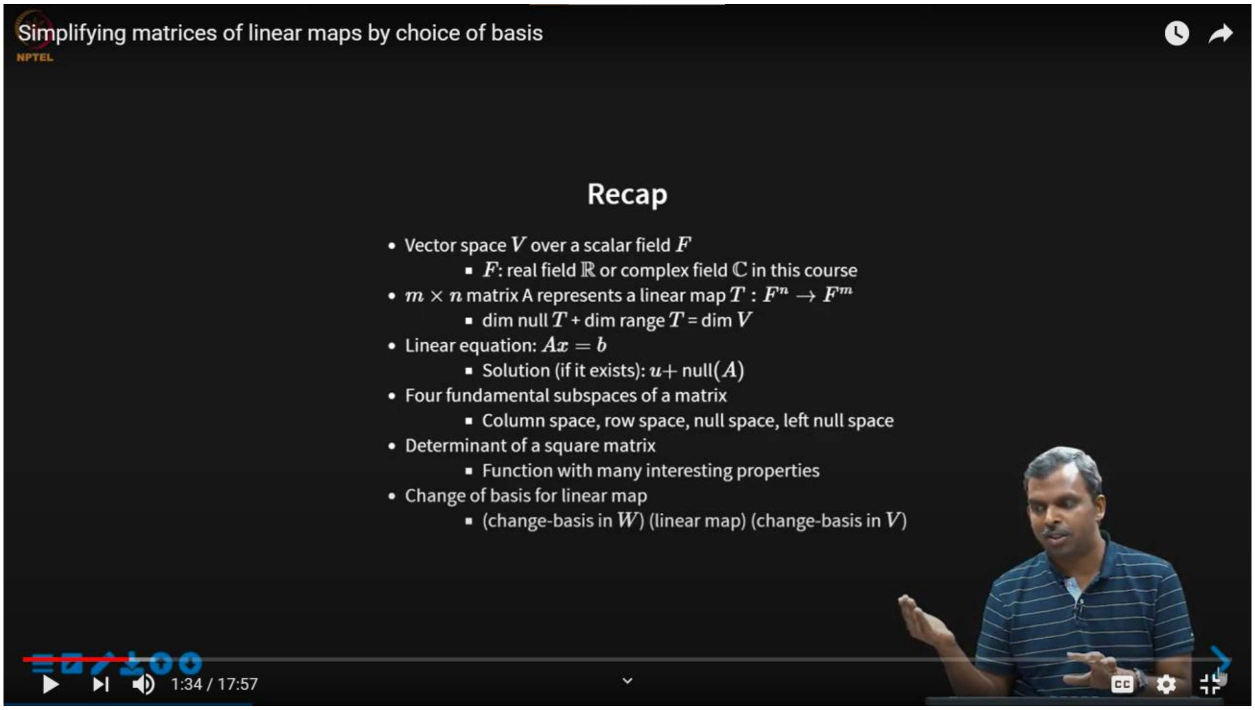
Applied Linear Algebra
Prof. Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras

Week 04
Simplifying matrices of linear maps by choice of basis

Okay, hello and welcome to this lecture. We're going to now talk about why change of basis or choice of bases is interesting for linear maps. It turns out the matrices representing linear maps change when the basis changes, okay? If you have a linear map given to you in some basis and it looks very complex, lot of entries, big matrix, maybe you can pick your basis in a clever fashion and get it to a very nice format, okay? So what is this nice matrix? Why are some matrices simpler than the others, let's talk about all that and what's possible when you change basis and what you can hope to achieve by changing basis for linear maps, let's look at that in this lecture.

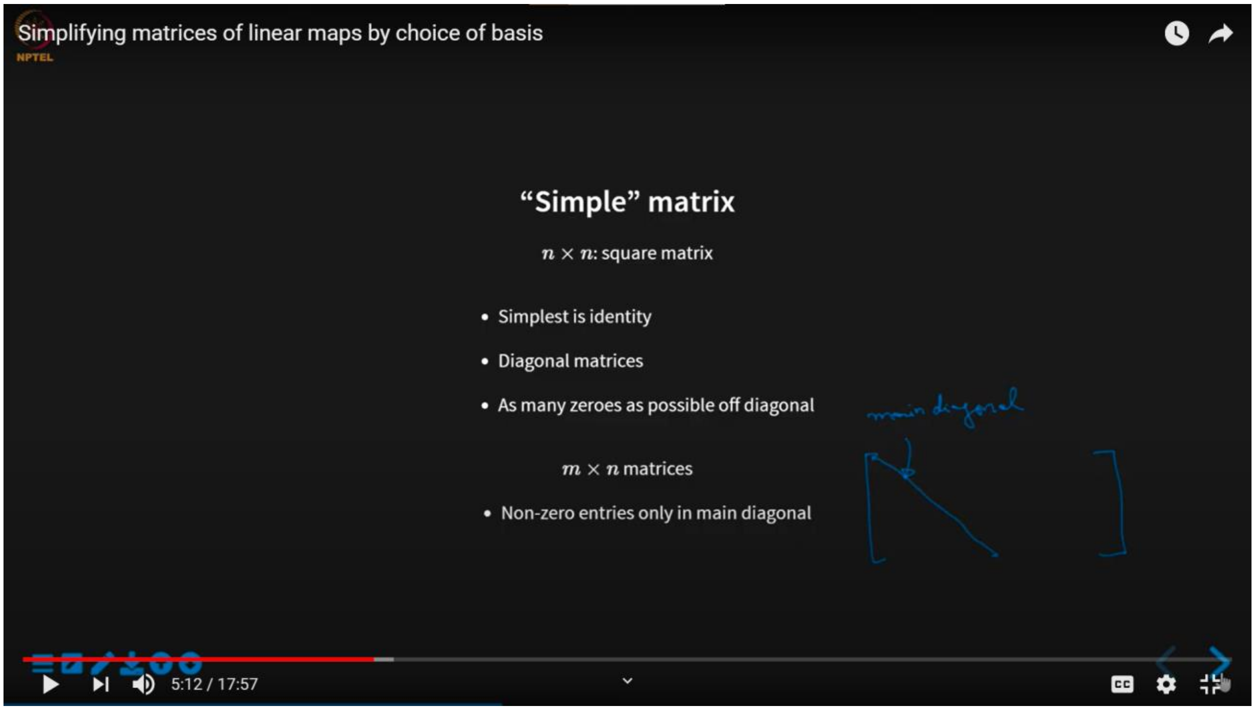
(Refer Slide Time: 01:34)



A quick recap. The most important part is the last point here which is change of basis for a linear map. The best way to think of it is: you have the linear map defined in terms of the original basis... So remember all the time when even though we think of the linear map in an abstract way, in most problems the linear map is going to be anyway specified with respect to one basis, okay? So always your starting point is the matrix specified in one basis, okay? So think of it like that, okay? So you have a matrix specified in one basis. How do you change bases? You look at the change basis coordinate matrix, right? So there is a way to do the identity map from one input basis to another. You do that on the first, okay? You go from the new basis to the old basis, you hit it with the old matrix, you get the old basis answer, then you transform to the new basis, you do change of basis, change of basis both ways and you can tackle any change of base situation. Put the suitable matrices in and you will get your answer, okay? So usually if it's V to W and V and W are different, you think of different bases. But quite often when it's an operator, you want to have the same basis and then you will have S^{-1} and all that and interesting operations coming, okay? So let us look at all those possibilities and why the choice of basis and change of basis gives you nice simplifications for the linear map, okay?

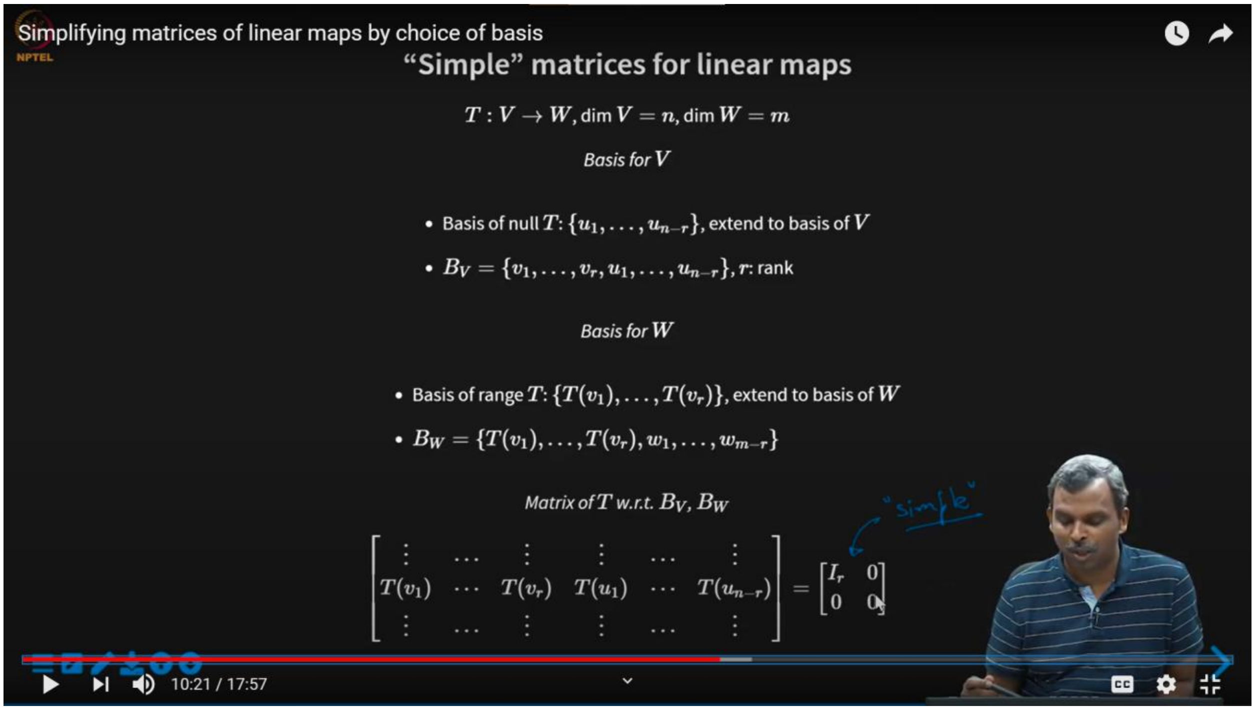
So first of all, what do we mean by a simple matrix? Why are some matrices simpler than the others? So if you think of $n \times n$ square matrices, all sorts of dimensions, you can have values all over the place, it can be quite complex in general, right? The simplest matrix you can think of is the identity matrix. Why? Because when it operates on any vector, it doesn't do anything, right? It gives you the same thing, okay? So you want to have a simple description to what happens to your input when this linear operator hits you, right? The identity is the simplest. Nothing happens, right? That's why it's simplest, okay? The next simplest presumably is diagonal matrix. Why is that? Because when, how do you describe what a diagonal matrix does to a vector? Each coordinate is scaled by the entry on the diagonal matrix, right? Corresponding entry on the diagonal matrix. So it's a very simple description for what happens to your input. So you can easily, you know, if in an actual system you can change the inputs, you know exactly what changing each input is going to do. You can easily predict, you can figure it out, okay? So that is very nice for diagonal matrices. So in general, the matrix becomes simpler as you have fewer and fewer non-zero elements in the off diagonal. Diagonal is very easy to describe. Anytime you have lots of non-zeros off diagonal, the matrix becomes more complicated to describe. Of course I mean you will take it with a pinch of salt. There are probably matrices with all entries non-zero but which are much simpler than having a few entries, you know? For instance, supposing you have a matrix with all ones, okay? So it's very easy to describe what the matrix does to the input, no? It just adds up everything and puts out the answer for every coordinate. It’s easy to describe. But still you know that's not the kind of example I'm looking at. I'm looking at a more complicated example when I say non-zero. So in general if you have many zeros then it is a good or simple matrix and maybe we have to aim for that. So maybe as you keep changing your basis you may cleverly do that so that you get to a form which is very nice for you, okay? So you remember when we did row operations, elementary row operations, column operations, we wanted to get to a very simple form, right? Again lots of zeros, it helps you deal with problems in a very nice way, okay? So what about $m \times n$ matrices? It's not too bad to extend it. So even if you have a non-square matrix, okay, even if you have a non-square matrix you can still think of the diagonal. So this is the main diagonal in a non-square matrix. Of course there will be many other diagonals of the same length but usually the first main diagonal is sort of thought of as a principal diagonal. So you want, you know, the main diagonal to have non-zero entries, maybe 1, you know or something like that. And then the other entries should all be 0 as much as possible, okay? So this is something that's nice to accomplish using change of basis. So let's see if this is possible, how much of this we can do. We already have this hint from this elementary row operations. So maybe it will help you. So let's see if all that is possible. But first we will look at it in a slightly abstract way, okay? So this is something that is easy to describe.

(Refer Slide Time: 05:12)



So you have a linear map $T: V \rightarrow W$, dimension of V is n, dimension of W is m. Let's begin with the basis for V which I will construct in a slightly clever way, okay? So I will take the basis of $null\ T$. First of all, I will take the rank to be r, okay? Since I am taking the rank to be r, the rank of, I mean the dimension of $null\ T$ is $n-r$, okay? That is the fundamental theorem. So you have $\{u_1, \dots, u_{n-r}\}$ as the basis of $null\ T$. Then I will extend this basis to a basis for V. But I will do this little bit of non-standard thing. I'll put the extension vectors first, okay? There's a good reason why I want to do it. So I'll put $v_1, \dots, v_r$ first and then $u_1, \dots, u_{n-r}$ next, okay? So I can always do this. This is, this will give me a basis for V. This will be the basis I pick. So notice once again what have I done. I find first the null space for T, okay? I mean this is not a very numerical procedure, but I am just talking about an abstract process here to tell you how to think of this, okay? So you find a basis for null and then you extend that to the basis for the vector space and put the extension vectors in the beginning. So that's what you do. You know the rank will be r. So the extension will have, you know, r guys in it. And then $u_1, u_2, \dots, u_{n-r}$. $n-r$ is the dimension for the null space, okay? So this is how it will be. Now what about basis for W, okay? You know this from one of the theorems. I mean the theorems for the fundamental theorem in fact works like this, no? The proof works like this, you pick a basis for the range of T. How do you pick the basis for the range of T? You know that $T(v_1), \dots, T(v_r)$ will form a basis for the range of T, right? So this comes from the fundamental theorem, from proof of the fundamental theorem. So you take that and extend it to a basis for W, okay? So I will have $T(v_1), \dots, T(v_r)$ and then I will add some additional vectors $w_1, \dots, w_{m-r}$ so that the whole thing becomes a basis for W. I've taken this $T(v_1), \dots, T(v_r)$ and extended it to form a basis for W. That's what I did, okay?

(Refer Slide Time: 10:21)

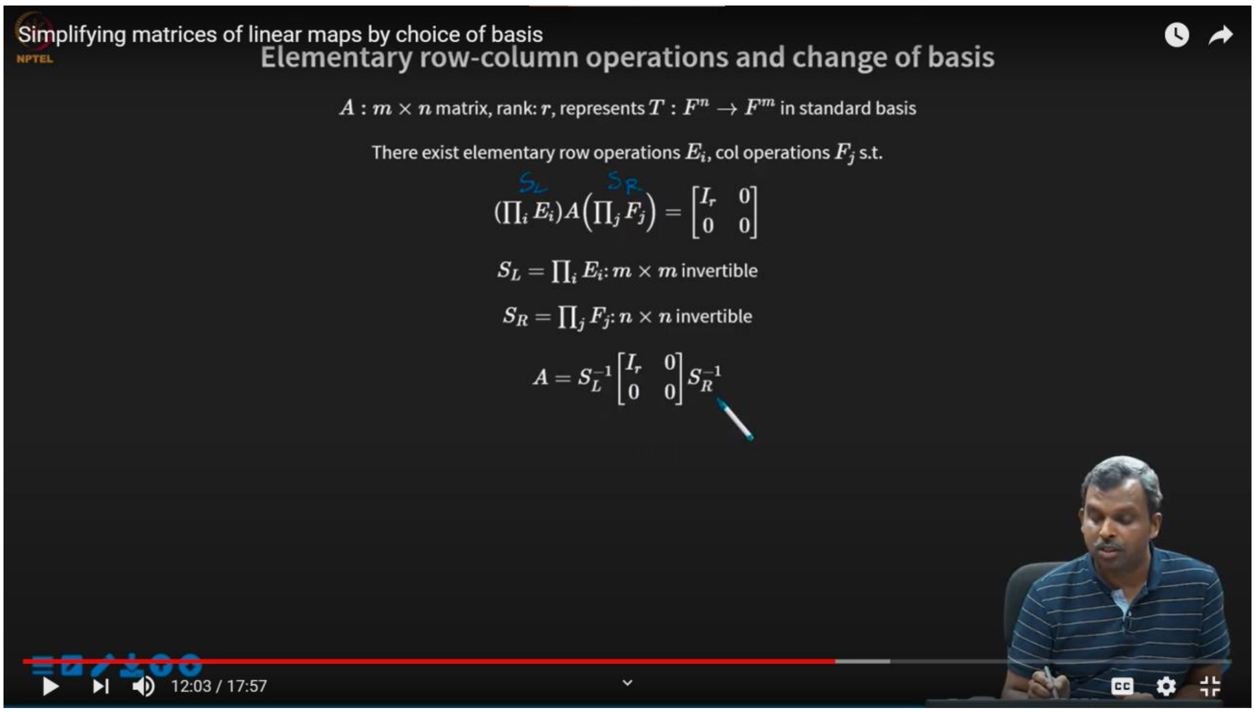


So this is how I will pick my basis B_v and basis B_w. These bases are very nice, you'll see what happens to the matrix shortly when you do this. But it's very nice. You can see why this is very nice. The reason is: if you look at the matrix of T with respect to B_v, B_w, what should I do? I should make a bunch of columns... The first column should be $T(v_1)$ represented in B_w, okay? I will come to that in a little while but the second one is $T(v_2)$. Likewise the r^{th} column is $T(v_r)$ and then I have $T(u_1), \dots, T(u_{n-r})$, okay? Now what happens to $T(u_1), \dots, T(u_{n-r})$? They are all 0, right? Because those are from the null, so they'll all go to 0. So the last $n - r$ columns will be all 0s here. What about the first r columns? $T(v_1)$ I have to represent it in this basis B_w. But that has $T(v_1)$ directly in the first position, so it will be 1 followed by a bunch of zeros. What will be $T(v_2)$? It will be 0 1 followed by zeros remaining. Up to $T(v_r)$ you will get this Ir in the first top left part, okay? So this basis, this particular way of choosing the basis B_v and B_w... By the way, it's not fixed, okay? So you can change any, I mean the basis for the null space can be different, the extension can be multiple extensions so you have so many different bases that give you this, B_v and B_w which give you this kind of a matrix finally. So this is easily, as you can imagine, is a very simple matrix, okay? So the point of doing these kind of abstract studies is just to tell you what's possible. So you know that you can pick a basis for V and W so that any linear map $T: V \to W$ has this very very simple form $[I_r\ 0;\ 0\ 0]$, okay? So it was easy to see, it was sort of hidden in the fundamental theorem in some sense. So that's why it's called the fundamental theorem, right? So there is very little new that you can do outside of it. So this is sort of hidden there. You pick your bases smartly motivated by the fundamental theorem, you will get a really, really simple matrix, okay? So this is nice to know.

Notice this even holds when W equals V. Don't think it doesn't hold for operators, okay? Even for operators you can do this. You can pick a basis B_v and a basis B_v', okay? So it cannot be equal, that is not guaranteed here, right? You pick a basis B_v and this B_w will not be the same as B_v just because W became equal to V. The way you are choosing it, you won't get the same thing, okay? So if you are allowed to pick two different bases, then you can get this very, very nice simple form. In fact the simplest form that's possible you can easily get, okay? So this is sort of like, you know, you can think of it as the fundamental theorem of representation of linear maps by matrices, you know? You pick your bases, you are allowed to pick whatever basis you want, this is the only linear map you have to worry about. Isn't it nice? Okay? So is this the end of everything? I mean we have nothing more to do. You've done, we've solved everything, right? But there is this problem with the basis B' being different from basis B_v, okay? So that's not very desirable always. Why is that? Why is it not very desirable?

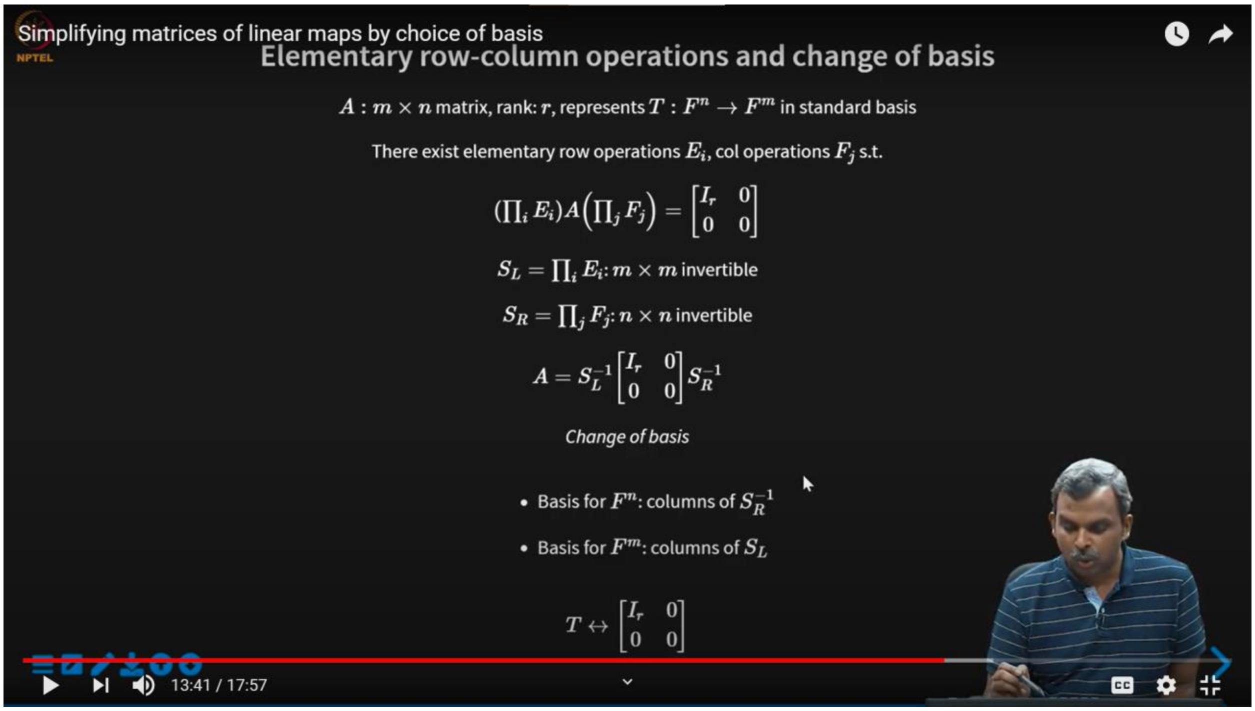
But before that, before that, before I look at why that's not very desirable, so let's look at these elementary row column operations. Remember this $[I_r\ 0;\ 0\ 0]$ also came from elementary row column operations, right? Now it turns out these elementary row column operations give you a precise recipe to find the basis B_v and B_w. So the previous slide also gave you some sort of a method. But that method is usually not very easy to implement. You have to find the null space, find a basis, find extension, maybe all that is not needed. Actually it's not needed. All you need to do is elementary row and column operations. You do that, you keep track of what matrices you had, you can directly find the B_v and B_w that gives you the simplest possible form, okay? Why is that? Take a look at this. We know that there exist elementary row operations, column operations E_i, F_j such that the product of E_i on the left product of F_j on the right with A gives you the simplest form. Now what else do you know? If I define S_l as the product of the row operation matrices, I get an $m \times m$ invertible matrix. I know this will be invertible. Similarly if I define my S_r as the product of all the right, the column operation matrices, I will get an $n \times n$ invertible matrix, okay? So if you look at this, you know I have S_l, right? So this is S_l. This is S_l, this is S_r. So if I want to push it to the right, I will get S_l^{-1}, S_r^{-1} on this side, right? So A becomes S_l^{-1}, this very simple matrix S_r^{-1}, okay? So that is something nice and what happens when you change the basis? So this actually sort of represents a change of basis, you know? See whenever you do change of bases, you're going to multiply it like this, like some sort of similarity with, you know, this and that be not being equal. S_l and S_r. So here is the basis that you pick for $\mathbb{F}^m$ and $\mathbb{F}^n$, okay? S_r^{-1} is here on this side. So I will pick the basis for $\mathbb{F}^n$ to be the columns of this S_r^{-1}, okay? Whatever S_r^{-1} I computed here, that, the columns of that is the basis for $\mathbb{F}^n$. And then the basis for $\mathbb{F}^m$, the output basis will be the columns of S_l itself, okay? If I do that, my linear map T becomes in this basis, in this particular basis it simply becomes the simple form $[I_r\ 0;\ 0\ 0]$, okay? So think about why that is true. It is, you can see it. But in any case if you are allowed to pick an arbitrary basis for $\mathbb{F}^n$ and an arbitrary basis for $\mathbb{F}^m$ and these two need not be equal, then you can pick them through this elementary row and column operation so that your, you know your linear map gets represented by just the $[I_r\ 0;\ 0\ 0]$. You do not have to do any null space basis extension and all that, just this linear elementary row column operation will give you what those bases have to be. In fact from here, you can also read off if you like the null space extension, range space and all of them, okay?

(Refer Slide Time: 12:03)



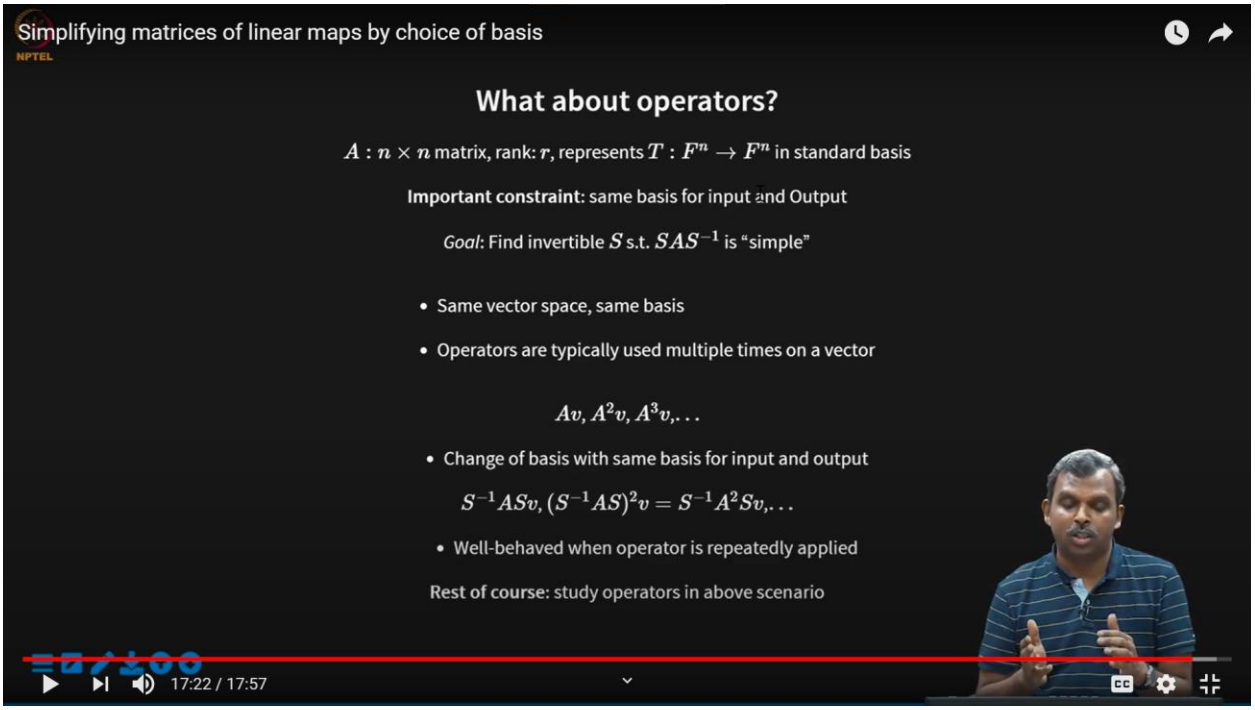
So this is sort of a summary now. So you see that any linear map has a very simple matrix representation on a suitable basis and at this point you have to really ask are we done. I mean, like, I mean how much simpler can we get than $[I_r\ 0;\ 0\ 0]$, what's the point in studying linear maps anything beyond this? Do we need to spend any more energy studying linear maps? Looks like it's very simple, right? But it turns out there is that little bit of problem with allowing, you know, different bases for V and W. When you go to operators in particular, there is a lot of motivation to keep the bases same for input and output, okay? So if you have particularly operators, $n \times n$ matrices, let's say rank is r. The transformation is from $\mathbb{F}^n$ to $\mathbb{F}^n$. Let's say in the standard basis you give a matrix. Usually this constraint becomes very important. You do not want one basis for the input, another basis for the output, okay? So in fact I want to find just one invertible matrix S such that SAS^{-1} under the similarity transform, SAS^{-1} I get a simple matrix. In the previous form that I did, I had S_l and S_r, two different matrices. That's not allowed, I want the same matrix, okay? So this makes life a little bit more complicated. It's not as easy as just arbitrary elementary row and column operations, I have to make sure they are balanced in some way, right? Whatever I do on this side I have to sort of undo on this side, okay? If you do that it's not so easy, okay? It becomes a little bit more tricky. It's not that easy, okay?

(Refer Slide Time: 13:41)



So why would you want to do this? Why do you insist on the same basis? It seems like same vector space, same basis it's reasonable, it's fair. But more than that, more than that, there is a really important utility for using the same basis. Notice what happens. When you have an operator, it's from the same vector space to the other and usually people will keep applying the operator multiple times. So these kind of operations are very, very useful in practice Av, you know, $A^2v, A^3v, \ldots, A^nv$, and we just keep on repeating the operator multiple times on the same vector v. So now notice what happens when you do a similarity transform. If you went to another representation for A keeping the basis for input and output the same, as you square, you get a very simple form, right? When you square it, this S and the next S^{-1} will cancel, okay? $(S^{-1}AS)^2$ will simply become S^1A^2S. It's just the same one basis transformation, and this A^2 gets preserved. If you had S_l and S_r not being inverses of each other, you won't have this cancellation. So you will have a very different looking, you know, linear map when you square it and that's not very nice, okay? The squaring not maintaining the linear map is a property that comes only from the similarity transform, only when you keep the basis for input and output the same. If you change it, you won't preserve this, okay? So and this behavior, you know, well behaved, a good behavior when you repeatedly use the operator is very important in practice, okay? So this is the, sort of the central reason why we continue to study linear operators in the rest of this course.

(Refer Slide Time: 17:22)



The first four weeks we've seen quite a few topics covering the basics of linear maps. Up to now we have looked at very, very fundamental basics of linear maps and matrix representation and fundamental theorem. What null space is, what range space is, how to simplify the matrix representation etc. etc. we have seen so far. It's very nice. Now for the next quite a few weeks, we will study operators in this above scenario, okay? So how do you fix, how do you do similarity transformations, how do you fix an input and output basis, transform from one basis to the other, how do you get very simple, you know, forms for the matrix, what is possible? It turns out what we did before is not possible, you can do only slightly lesser and this also leads naturally to the study of very important concepts like eigenvectors, eigenspaces, what they all mean, okay? So this is sort of like a critical point for the course. End of week four. We are done with the foundational aspects of, you know, vector spaces, you know, linear maps and all that. We will start diving into eigenvalues and eigenvectors and eigenspaces from the next week onwards, okay? Thank you very much. Hope you're having fun. Hope you enjoy the rest of the course also. Thanks.