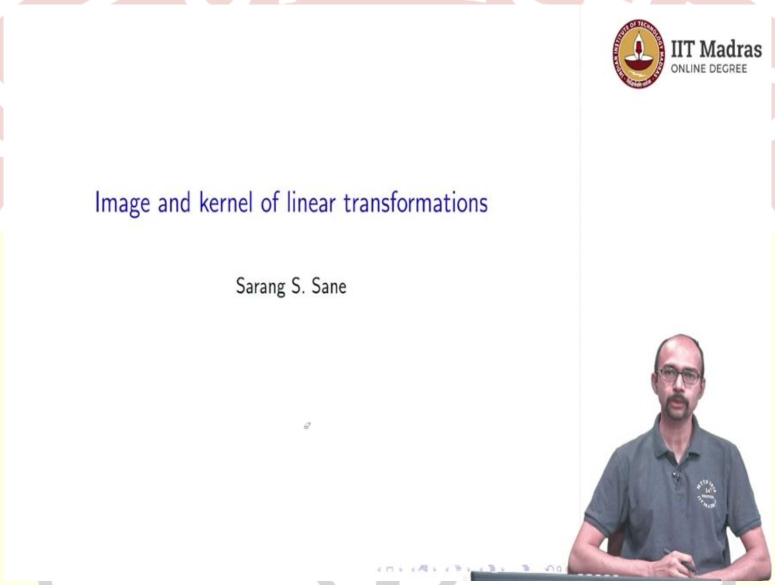
IIT Madras

ONLINE DEGREE

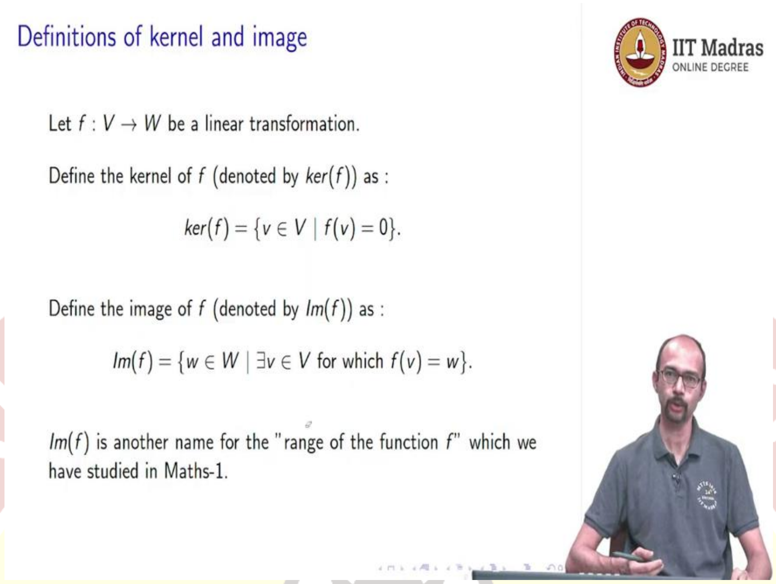
Mathematics for Data Science - 2
Professor. Sarang S. Sane
Department of Mathematics
Indian Institute of Technology, Madras
Image and kernel of linear transformations

(Refer Slide Time: 0:14)



Hello, and welcome to the Maths 2 component of the online B.Sc. program on data science. In this video we are going to study the notion of the image and kernel of linear transformations. So, remember previously that we have studied the notion of the null space, so this is some, for a matrix. So, this is some analogues thing for a linear transformation. So, for a matrix we know what is the row space or the column space and that is, the analogous notion for a linear transformation is the image.

(Refer Slide Time: 0:44)



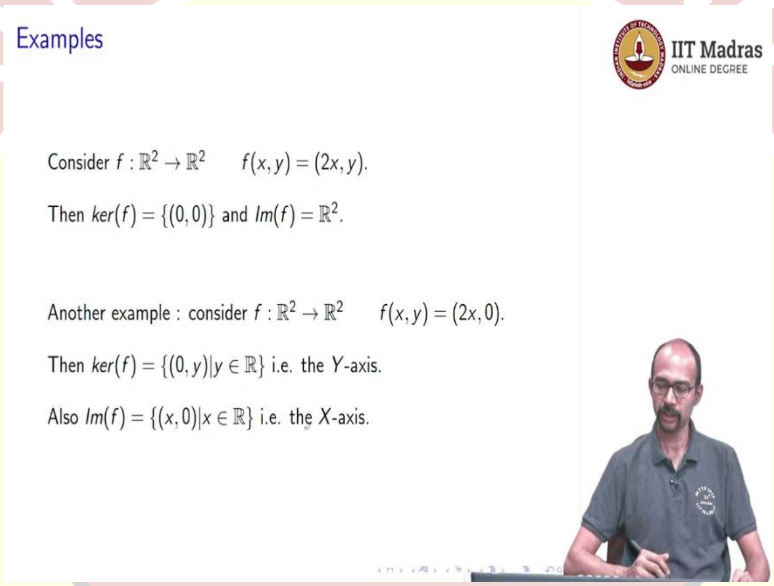
So, let us start by, from the definitions. So, let $f: V \to W$ be a linear transformation and define the kernel of f, which is denoted by $ker(f)$ as kernel of f is all those vectors v, such that $f(v) = 0$. So, V and W are vector spaces, f is a linear transformation and we are looking at all those vectors so that $f(v) = 0$. Now, it is easy to check that kernel of v, a kernel of f is actually a subspace of V, which means it is a vector space in its own right. What that means is if you take a linear combination of elements of kernel of v, they still belong to kernel of v.

Define the image of f. This is denoted by $Im(f)$. Sometimes we will use a little i and sometimes we will use capital I. So, that may, that notation will be a bit vague. So, we will define it as all those vectors w. So, that there exists v in V. So, that $w = f(v)$. So, this is nothing but the usual definition of the range of a function. If you have a function between two spaces or two sets, we call the range of f as the set of values which the function takes. So, that this is exactly the range of the function. So, image of f is another name for the range of the function f. And I think you have studied range of a function in Maths 1.

In case you have not, I just recall for you that range of a function is the set of values which the function takes. So, anyway, if you did not follow what I just said, we will anyway look at some examples. Before that, let me also comment on the previous slide that the image of f is also a subspace of W. So, if two vectors are in the image, the sum is also in the image, because you can take, so if $w_1 = f(v_1)$ and $w_1 = f(v_2)$, then you can look at $f(v_1 + v_2)$that is exactly $w_1 + w_2$ because remember f is a linear transformation and that means $w_1 + w_2$ is also in the image.

Similarly, if you have αw, where w is in the image, then w is $f(v)$. So, you look at αv, then $f(\alpha v) = \alpha f(v) = \alpha w$. So, αw is also in the image. Both of these remember use, I mean, in the proofs that I gave, the statements that I just made, the fact that it is a linear transformation was crucially used. So, this is not true for functions and so on for other things.

(Refer Slide Time: 3:25)



So, let us look at some examples to get our idea straight. So, look at the function from $\mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (2x, y)$. First of all, why is this a linear transformation, because each coordinate is a linear combination. Remember that linear combinations, if each coordinate is a linear combination that is exactly what a linear transformation is when it comes to functions from $\mathbb{R}^n \to \mathbb{R}^m$. So, $2x$ is a linear transformation of x and y, and y is a linear transformation of x and y. So, this is indeed a linear transformation. So, each of these are linear combinations. So, this is a linear transformation.

So, what is the kernel of f? The kernel of f is all those values so that f of that number x, that vector (x, y) is 0. 0 where, 0 in the codomain $\mathbb{R}^2$. So, 0 in the vector space $\mathbb{R}^2$. So, 0 in $\mathbb{R}^2$ is (0, 0). So, in the previous slide also note that we said $f(v) = 0$. By that 0 we mean the 0 in W. So, here we have, we want to look for those x, y such that $f(x, y) = (0, 0)$. So, it is clear that if

$(2x, y) = (0, 0)$, then $2x$ is 0 and y is 0, which will tell us that x is 0 and y is 0. So, the kernel is exactly $(0, 0)$. So, the kernel is the 0 element or the 0 vector in $\mathbb{R}^2$.

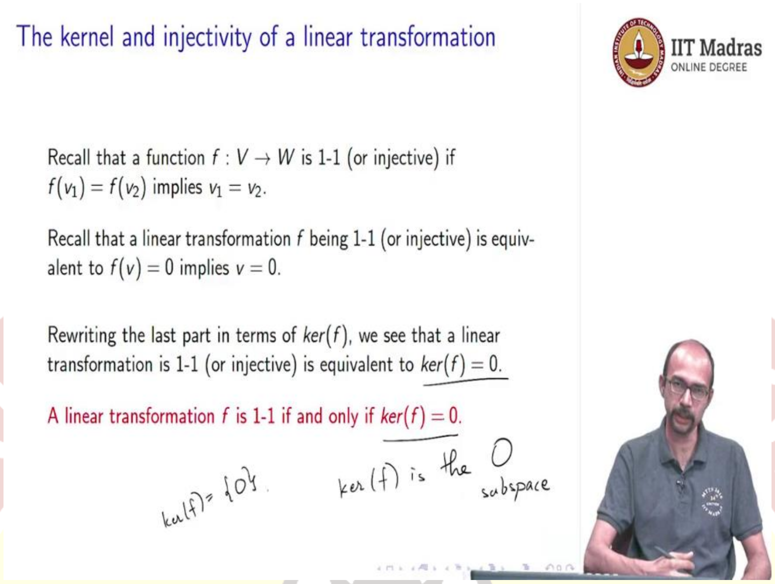
What is the image of f? The image of f is the entire $\mathbb{R}^2$. Why is that? So, if you want to write a vector W in $\mathbb{R}^2$ as the image of some element, then what would you do? You would look at, suppose $w = (u, v)$, then you look at $(\frac{u}{2}, v)$ And if you apply f on $(\frac{u}{2}, v)$, you get $(2 * \frac{u}{2})$, which is u and v. So, that tells us that u comma v is in the image. So, every vector u comma v is in the image that means the image is the entire space $\mathbb{R}^2$, the entire vector space.

Consider the example, $\mathbb{R}^2 \to \mathbb{R}^2, f(x, y) = (2x, 0)$. Again, this is a linear transformation because each coordinate is a linear combination. So, what is the image and what is the kernel? So, kernel of f is $(0, y)$. Why is this? Suppose $f(x, y) = 0$, we want to, from there we want to solve for (x, y). So, that means $(2x, 0)$ is equal to $(0, 0)$. That means $2x$ is 0 and 0 is equal to 0. So, 0, 0 does not yield us anything on (x, y), any restriction on (x, y). $2x = 0$ yields us the restriction $x = 0$, that means y can be anything. So, the kernel of f is $(0, y)$. So, this is the entire y-axis. $(0, y)$. is the y-axis.

And what is the image of f? Image of f is, well, you have $f(x, y) = (2x, 0)$. So, now suppose you look for an element (u, v) and you ask when can $(u, v) = (2x, 0)$. So, the first thing you need is that v is 0. That means this element must, this vector must lie on the x-axis. And then you have to ask whether on the x-axis it can take all possible values.

So, you get the equation u is $2x$, which if you solve, you get x is $\frac{u}{2}$, which means if you put $(x, y) = (\frac{u}{2}, 0)$, then the image of that element $f(x, y) = (u, 0)$. So, you get that every element, every vector on the x $-$axis is indeed in the image. And you already saw that the image is contained in the x $-$axis that means the image is equal to the x-axis.

(Refer Slide Time: 7:40)



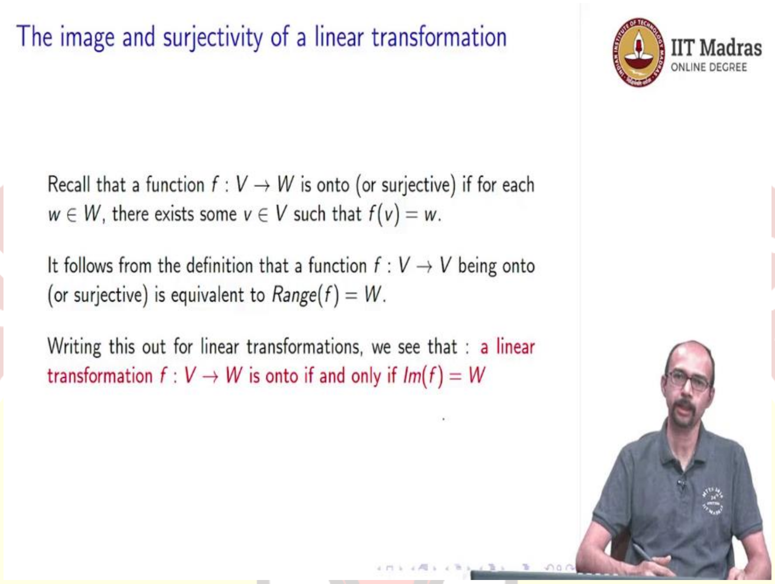
So, let us go on to studying the relation between the kernel and injectivity, which, so we remember that a function is injective or one-one, if $f(v_1) = f(v_2) \to v_1 = v_2$ and for linear transformations this is the same as demanding that $f(v_0) \to v = 0$. We prove this actually. So, what does this have to do with the kernel? Well, the first one may, you may not be, it may not be clear, but the second statement makes it clear what it has to do with the kernel. $f(v = 0) \to v = 0$

So, if a linear transformation is one-one or injective then the kernel must be 0. That is what we are saying, because if $f(v) = 0$, then $v = 0$. So, the only possible choices of vectors which yield that if $f(v) = 0$ is the 0 vector itself. So, if f is one-one or injective, then the kernel of f is 0. On the other hand, if the kernel is 0, that means whenever if $f(v) = 0$, v is 0, so that means the condition for being one-one is satisfied. So, kernel of f is 0 also means that the linear transformation is one-one.

So, being the linear transformation being one-one is equivalent to kernel of f is 0. That is what we have proved here. So, linear transformation f is one-one if and only if kernel of f is 0. So, what do I mean by kernel of f is 0. What I mean here is that it is a 0 subspace. So, this should be better written as kernel of f is the 0 subspace, which means it consists of the set 0. And remember, 0 itself is a vector space. It satisfies all the properties of a vector space. And it is a subset of V.

So, it is a subspace of V. So, kernel of f is the 0 subspace. That is what we mean by these two statements here. Kernel of f is 0 and kernel of f is 0. So, be careful when we write this. The 0 in this case here on the right is a 0 subspace and not the 0 vector. It means it is a set consisting of the 0 vector.

(Refer Slide Time: 10:08)

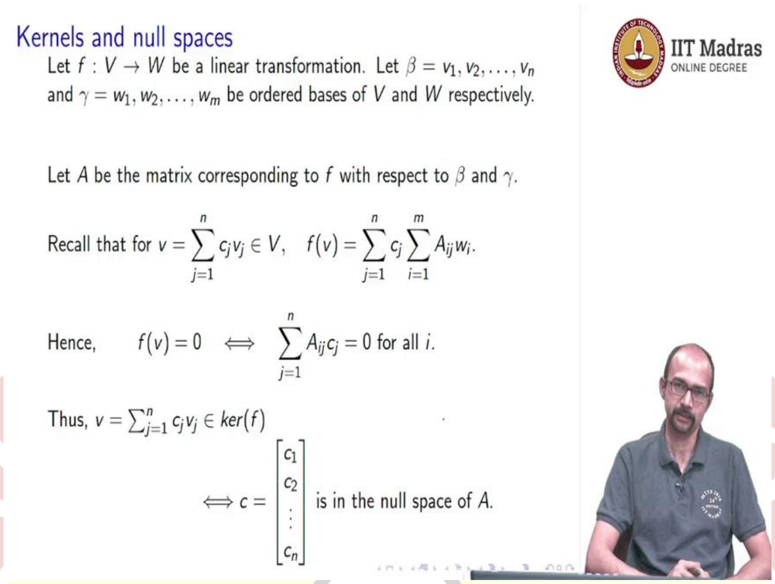


So, let us again see the relation like kernel is related to being injective let us study what is the relation between the image and the subjectivity of a linear transformation. So, recall that a function $f: V \to W$ is onto or surjective if for each w in W there is some v in V such that $f(v) = w$. In other words, the range of this function is the entire W. So, this is actually, this part is the same for any function. This is not, this has nothing to do with linear transformations per se, what the statement in this slide.

So, it follows from the definition that a function f being onto is equal to range of f being W that is exactly what we just commented. So, for linear transformations, in particular, this means that because remember range is the same as image, only thing is when we say range we think of it as a set and when we say image we think of it as a subspace. This is the main difference. Otherwise, there is no real difference between the range and image as such.

So, a linear transformation is onto if and only the image of f is W. And this image of f is W is being concluded as a subspace. Image of f is actually a subspace of W always and we say that it is onto. We have proved that it is onto, is equivalent to the subspace image of f actually been the entire space. But this is not very different from what happened for functions.

(Refer Slide Time: 11:57)



Let us now proceed and make the connection between kernels and null spaces. So, so far there is no matrix in the picture. So, now, we are going to introduce some matrix in the picture and we will associate the kernel of this linear transformation with that matrix. So, let $f: V \to W$ be a linear transformation, let β be $v_1, v_2, \dots, v_n$ and γ be $w_1, w_2, \dots, w_n$ which are both, both of which are ordered basis. β is an ordered basis of V and γ is an ordered basis for W.

So, now, indeed we can talk about a matrix, because remember in the previous video, we have talked about how we, if you have a linear transformations along with an ordered basis for each of V and W, there is a corresponding matrix. And how do we write down that matrix? That write, we write down that matrix by looking at $f(v_i)$ and expressing that in terms of $w_j{'}$s, as a linear combination of the $w_j{'}$s and whatever coefficients we get go into the matrix. So, the coefficient of f, coefficients of $f(v_i)$ go into the ith column of that matrix.

So, let A be the corresponding matrix. So, just recall that for v in, so suppose $v \sum c_j v_j$ then $f(v) = \sum c_j, f(v_j)$ that is by the linear transformation property, axioms of being a linear transformation. And then, f of, so summation $f(v_j)$ can be further written as $\sum A_{ij} w_i$. So, $f(v_j) = \sum A_{ij} w_i$. So, that is corresponding the jth column of the matrix A, that is what these, where these coefficients are coming from. That is exactly what we just commented.

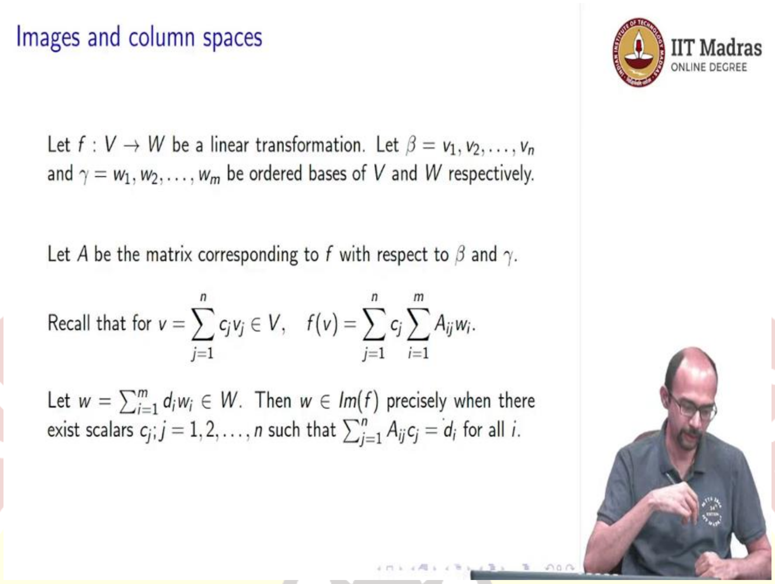
So, when is $f(v) = 0$? So, $f(v) = 0$ if and only if each coefficient in this linear combination of $w_i{'}$s is 0. So, what is the coefficient of w_i? So, you have to take that summation i is equal 1 to

m outside and take the summation j is equal to 1 to n inside and then you will get $\sum \mathrm{A_{ij} c_j}$ over j that is the coefficient of w_i. And then we know that if you have a linear combination of w_i's which is 0 then each coefficient must be 0. Why is that, because wis are a basis, remember, and basis satisfies that it is linearly independent.

So, if summation of some linear combination, this particular linear combination of w_i's is 0 then each coefficient is 0. So, $f(v = 0)$ is exactly the same as saying that $\sum \mathrm{A_{ij} c_j} = 0$. And now this equation should ring a bell in your mind. So, the left hand side is saying, so $f(v) = 0$ means this is exactly the definition of v being in the kernel of f. And the right hand side is saying that if you look at the vector c consisting of $c_1, c_2, \ldots, c_n$, then that vector is in the null space of A.

So, the relation we have between the kernel and the null space is the following. That if you have a vector v, which when you express as a linear combination of the v_j's, there, remember, it is an ordered basis, so you can write it as $\sum \mathrm{c_j v_j}$in a unique way. So, this vector is in the kernel if and only if this happens precisely when if you write these coefficients of the v_j's as a column, then this column must be in the null space of the matrix A. So, that is the relation between kernels and null spaces.

(Refer Slide Time: 16:22)

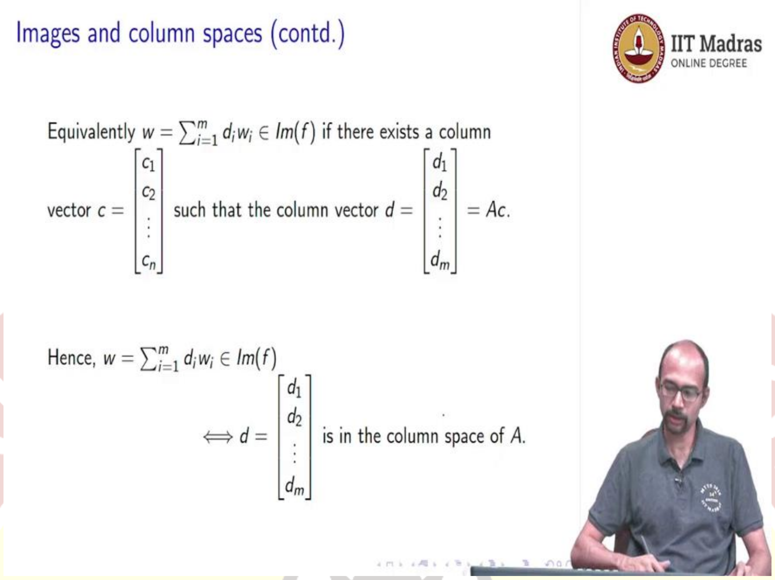


So, let us similarly describe what is the relation between images and column spaces? So, let V be a, $f: V \to W$ be a linear transformation, let β be $v_1, v_2, \dots, v_n$, γ be w_1, w_2, w_n both of these are ordered basis, β for V and γ for W and let A be the matrix corresponding to f with respect to β and γ. So, we know how to write this. The jth column of A corresponds to the coefficients arising as expressing $f(v_j)$ in terms of w_i's.

So, now, again recall that for v is summation v_j, $c_j v_j$ in V, $f(v) = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} A_{ij} w_i$, which if you take the summations, interchange the summations we can do that because it is a finite sum then you get that this is summation i is 1 through m, summation c_j, $\sum A_{ij} c_j * w_i$. So, the coefficient of w_i is $\sum A_{ij} c_j$ over j. So, suppose now w is $\sum d_i w_i$ and it is, so this is a vector in W. So, these dis are unique, because wis form of basis. So, you can write w_i's $\sum d_i w_i$. So, when does this lie in the image.

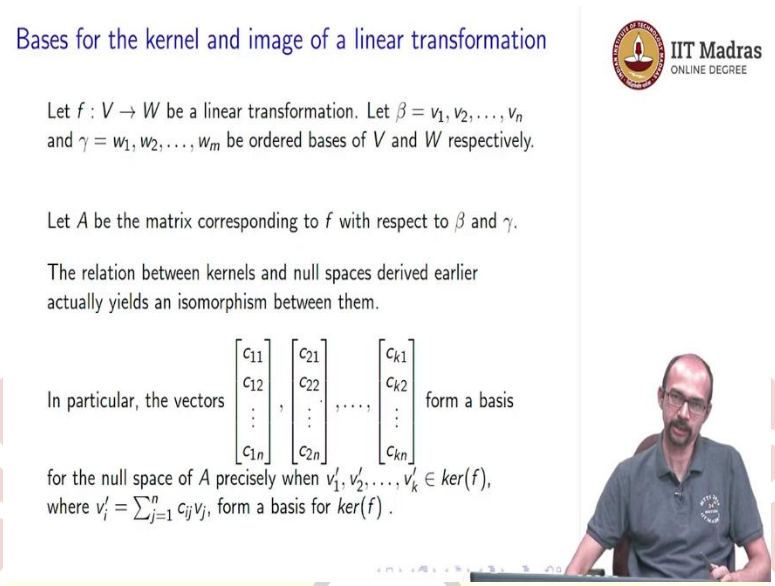
So, this lies in the image precisely when there exists a scalar c_j so that $\sum A_{ij} c_j$ is d_i. So, the scalars from the coefficients of something in the image, something of the form $f(v) = \sum A_{ij} c_j$. So, that means these d_i should be $\sum A_{ij} c_j$ in order for w to be $f(v)$, where, what is v? V is equal to $\sum c_j v_j$.

(Refer Slide Time: 18:34)



So, what that means is, we can now express this in terms of columns. What it means is that $\sum d_i w_i$ is in image of f if there exists a column vector c such that the column vector $d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix} = Ac$. So, but remember Ac is exactly a linear combination of the columns of A. So, what that means is that w is in the image of f where w is given by $\sum d_i w_i$ if and only if you look at the vector $d = d_1, d_2, \ldots, d_m$ and this is in the column space of A. So, that is a relation between the image of f and the column space of this matrix A.

(Refer Slide Time: 19:26)



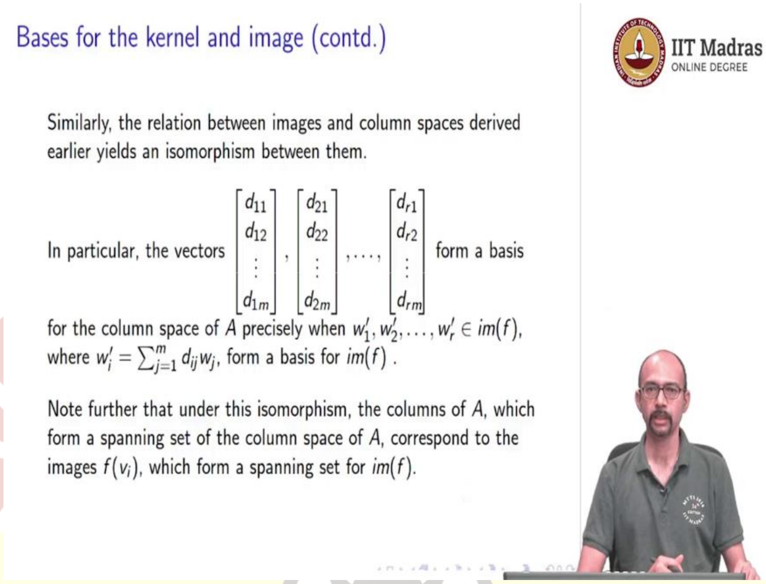
Now, let us see if we can use this to find the basis of the kernel and the image of a linear transformation. So, again let $f: V \to W$ be a linear transformation, choose ordered basis β and γ for V and W, respectively. Let A be the corresponding matrix. So, the point, I mean, the first point here is that we can look at the kernel and the null space and what we have really shown is that the kernel is isomorphic to the null space. That is what really we showed in the slide on kernels and null spaces.

Why is that? That is because for every element of the kernel, we have a unique element of the null space. So, you can define a function and it is easy to verify that this is indeed a linear transformation from the kernel to the null space. And it is also equally clear that this function is both one-one and onto that something you can check. So, that tells us that the kernel is actually isomorphic to the null space.

So, in particular, if you look at the vectors $\begin{bmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1n} \end{bmatrix}, \begin{bmatrix} c_{21} \\ c_{22} \\ \vdots \\ c_{2n} \end{bmatrix}, \dots, \begin{bmatrix} c_{k1} \\ c_{k2} \\ \vdots \\ c_{kn} \end{bmatrix}$ these form a basis for the null space of A precisely when v_1, the corresponding v'_1, v'_2, …,v'_k lie in kernel of f, form a basis for kernel of f. And what is the relation between v_i's and, v'_i and these columns? The coefficients of $v'_i$when expressed in terms of v_j are exactly corresponding to the jth, the ith vector in this collection.

So, in other words, what is the main point? The main point is I can read off a basis for the kernel if I am, if I know how to get a basis for the null space. And indeed, we do know how to get a basis for the null space.

(Refer Slide Time: 21:46)



Similar to the previous situation, we can find the relation between images and column spaces specifically about basis between them using the isomorphism that we have discussed earlier. So, if we consider the column vectors $\begin{bmatrix} d_{11} \\ d_{12} \\ \vdots \\ d_{1m} \end{bmatrix}, \begin{bmatrix} d_{21} \\ d_{22} \\ \vdots \\ d_{2m} \end{bmatrix}, \ldots, \begin{bmatrix} d_{r1} \\ d_{r2} \\ \vdots \\ d_{rm} \end{bmatrix}$ then these form a basis for the column space of A precisely when if you look at the corresponding linear combinations of the $w_i{}'$s, w_i' is $\sum_{j=1}^{m} d_{ij} w_j$ these form a basis for image of f. So, this follows because of the previous isomorphism that was discussed.

So, once we have this isomorphism note also that if you look at the columns of A which form a spanning set of the column space, then these correspond to the images of $f(v_i)$. So, the images f of v_i which form a spanning set for image of f. So, what is the point of all this? So, as we have noted before, we can convert the problem of finding a basis for the image space to the problem of finding a basis for the column space.

And, well, that is something we know how to do, because to find a basis of the column space, what do we do, we look at the columns, then we do row reduction and once it is in row, reduced row echelon form, you look at the columns containing the pivot elements, and then the original columns, those are the ones which form a basis for the column space. So, now, what this is telling us is, once we know those columns which form a basis for the column space, you look at the corresponding linear combinations of the $w_j's$ and that will give you a basis for the image of f.

So, this video, we have seen the notion of image and kernel and we have discussed these relationships between the kernel and the null space of the associated matrix once we fix basis, ordered basis, and similarly, the relationship between the image of a linear transformation and the column space of a matrix which we obtain by fixing ordered basis.

So, we have seen that these are isomorphic. You know the interplay between these two. And using these we can get basis by considering these corresponding matrices that we obtain and then obtaining basis for those using row reduction, which we know how to do from the last week. So, in the next video, we will implement all of this in concrete examples. Thank you.