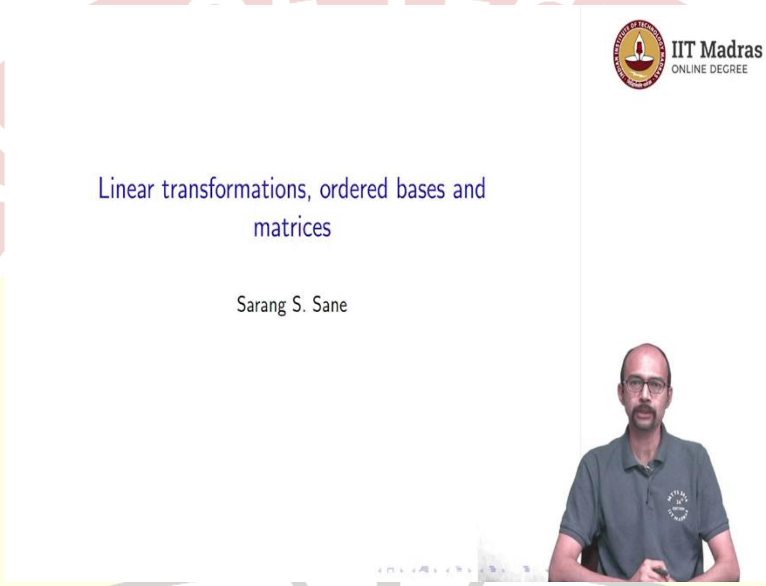
IIT Madras

ONLINE DEGREE

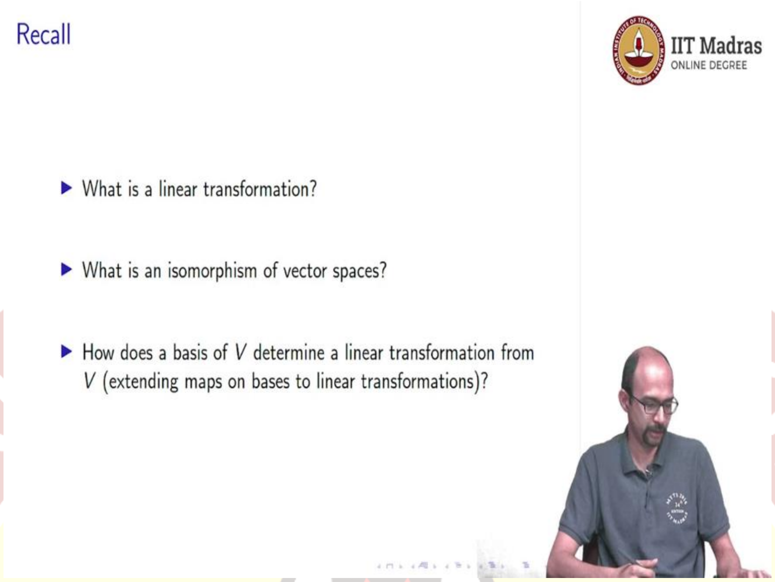
Mathematics for Data Science - 2
Professor Sarang S. Sane
Department of Mathematics
Indian Institute of Technology, Madras
Linear transformations, ordered bases and matrices

(Refer Slide Time: 00:14)



Hello, and welcome to the Maths 2 component of the online degree on Data Science. In this video, we are going to talk about linear transformations and their relation with ordered bases and matrices. So, this is continuing on our previous video where we have defined linear transformations. We have already seen the connection, a little bit of connection between bases and matrices with linear transformations and we are going to expand on that in this video.

(Refer Slide Time: 00:41)

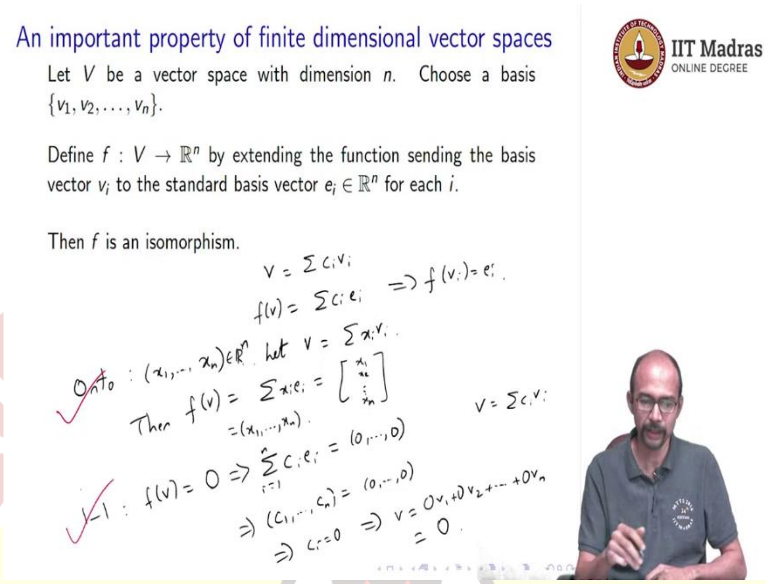


So, let us recall first what we have done so far. We have studied what is a linear transformation. So, a linear transformation was a function between two vector spaces. So, a function f from v to w, let us say, which satisfies linearity. So, linearity is the condition that $f(v_1 + cv_2)$ is $f(v_1) + cf(v_2)$. Equivalently, we could write it as two separate conditions $f(v_1 + v_2) = f(v_1) + f(v_2)$, and $f(cv_1) = cf(v_1)$. And you can use this to say that $f(\sum(c_i v_i)) = \sum c_i f(v_i)$. In other words, sum comes out and the scalars come out. That is a linear transformation.

We have seen what is an isomorphism of vector spaces, namely if you have a linear transformation which is a bijection that is what we called an isomorphism of vector spaces. And finally, we saw in our previous video that a linear transformation is determined by the values it takes on bases vectors.

Equivalently or other conversely, if you have a set of values, a set of vectors in the codomain w, and you define a function on the bases given by the $f(v_i) = w_i$, you can extend this function to a linear transformation f v to w and we explicitly wrote down what, how to extend it. So, extending maps on bases to linear transformations.

(Refer Slide Time: 02:21)



So, let us continue from this point and study first very important property of finite dimensional vector spaces. So, let V be a vector space with dimension n. Choose a bases v_1, v_2, v_n. And we are going to define a function f from V to $\mathbb{R}^n$, which sends v_i to e_i. So, how do we do that? Well, we define the map $f(v_i) = e_i$. This is only on the basis v_1, v_2, v_n. So, it is a map on the set v_1, v_2, v_n. And then we extend it to a linear transformation, which by abuse of notation, we are going to still call f. And, of course, as a result, it continues to take v_i to e_i. So, the claim is that v is an iso, sorry f is an isomorphism. Let us check why that is the case.

So, first, let us write down what is the function. So, if you have v is $\sum c_i v_i$, then $f(v) = \sum c_i e_i$. And we saw as soon as you do this that from here we obtained that $f(v_i) = e_i$. Now, why is this a isomorphism? We have seen already that this is the linear transformation given by extension. So, it is a linear transformation. I did not check that in the previous video, but I hope that you have checked it. If not, you should do it right now when you finish this video.

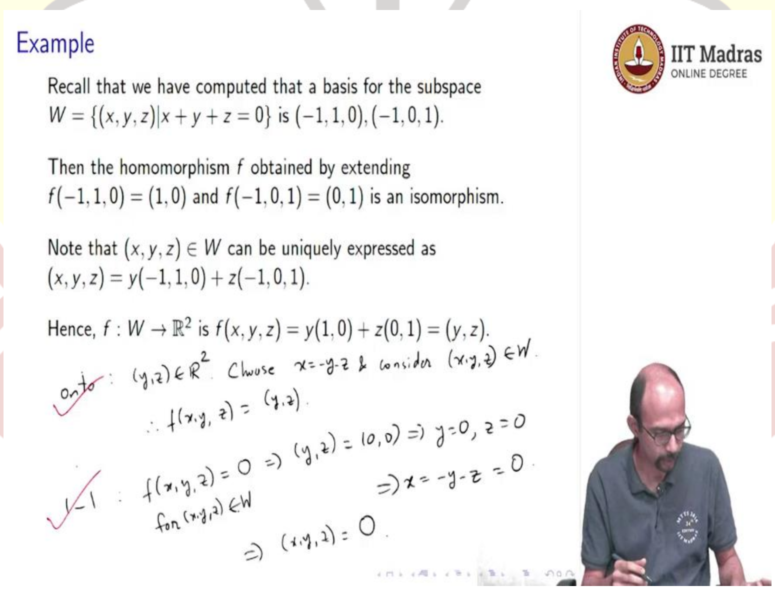
So, $f(v) = \sum c_i e_i$ and $f(v_i) = e_i$. So, let us check that is an isomorphism. To check that, we have to check it is one-one and onto. So how do we check onto? So, onto, how do we check that? Well, let us take a element of $\mathbb{R}^n$. Let us say, I have x_1, x_2, x_n. Let c_i, let $v = \sum x_i v_i$, so the c_i is x_i. So, this is in $\mathbb{R}^n$. Then what is $f(v)$? $f(v) = \sum x_i e_i$, which is exactly, in terms of, as a column vector, this is the column vector x_1, x_2, x_n. And here, when we say we are going to write it as a row vector

$f(v)$ is the row vector corresponding to this, so $f(v)$ is exactly x_1, x_2, x_n. So, this means it is onto. So, therefore it is onto. I have checked that it is onto.

Now, let us check one-one. It is a linear transformation. So, it is enough for me to check that $f(v) = 0 \rightarrow v = 0$. But what does it mean to say that $f(v) = 0$? First of all, what is 0 in $\mathbb{R}^n$, it is the origin or the vector (0, 0, 0). So, let us put in the definition of $f(v)$, which is written above. So, $f(v) = \sum c_i e_i$, this is 0, 0, 0. But what is this $\sum c_i e_i$? So, v is written as $\sum c_i v_i$. These are those, the $c_i{}'$s are those coefficients coming out of the unique way of writing $v = \sum c_i v_i$, where i of course runs 1 through n.

But on the other hand, this is the column vector (c_1, c_2, c_n). As we have been consistently doing, I am abusing notation between the rows and columns because there is no matrix multiplication. So, in other words, this isc_1, c_2, c_n but this is (0, 0, 0). That means each of the $c_i{}'$s are 0. But what is that so for all i, that means $v = 0v_1 + 0v_2 + 0v_n$. So, these are all the scalar 0, which evaluates the vector 0 in your vector space v. So, I have proved that it is one-one. That is why it is an isomorphism.

(Refer Slide Time: 07:00)



So, last video, we did not see any examples of linear transformations which are not linear mappings. All of them were on $\mathbb{R}^n$ and $\mathbb{R}^m$ and so on. So, this is our first example where we are going to study a linear transformation. So, let us look at the subspace W of $\mathbb{R}^3$, which is given by

x, y, z, where $x + y + z = 0$. We have checked in a previous video that this is indeed a subspace. This was the null space of something. I will leave that part to you to recall what, where we did this earlier.

So, this is a subspace of $\mathbb{R}^3$, so which means in its vector space in its own rate, by which we mean that if you take two elements of the form x, y, z such that $x + y + z = 0$, then they add up and you get a new vector which is also satisfying the same relation that $x + y + z = 0$. And if you take a scalar multiple, the same thing happens. So, we had worked out earlier that a basis for this subspace is $(-1, 1, 0)$ and $(-1, 0, 1)$. I will, you should check this.

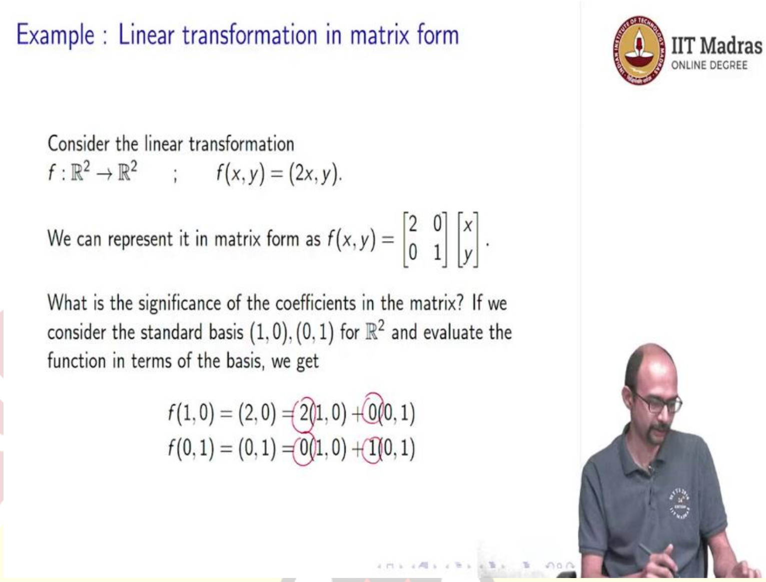
So, the homomorphism f obtained by extending $f(-1, 1, 0)$ is $(1, 0)$ and $f(-1, 0, 1)$, is $(0, 1)$ is an isomorphism. Why is that? That is because if you write any x, y, z in W, it can be uniquely expressed as $y\,(-1, 1, 0) + z(-1, 0, 1)$. You might wonder why that is the case, because you get $-y, -z$ in the first coordinate, and then y in the second and z in the third coordinate, but $-y - z = x$. Why is that, because x, y, z is in W, that means $x + y + z = 0$ that means $x = -y - z$ So, this is fine. And it is unique because this is a basis.

So, then we know how to extend functions, which are defined on the basis to linear transformations. So, you get that $f(x, y, z) = \ y\,(1, 0) + z\,(0, 1) = (y, z)$. So, this is actually the projection to the YZ plane. This is a projection from w onto the YZ plane. I did not check why it is an isomorphism. But it is clear it is an isomorphism. I will claim because let us see onto, so projection map, it is always going to be onto.

So, if you want y, z in $\mathbb{R}^2$ so choose $x = -y - z$ and consider x, y, z in W. It does lie in W, because $x + y + z = 0$. So, therefore, $f(x, y, z) = y + z = (y, z)$. So you have to choose x, otherwise we do not know x, y, z is in W. So, you have to get something which is in W. So, therefore, it is onto. So, we have checked onto.

And one-one is very direct. So, if $f(x, y, z) = 0$, so what is 0 in $\mathbb{R}^2$, it is (0,0) and let us put in the value of $f(x, y, z) = (0, 0)$. So, this means $y = 0$ and $z = 0$. But that means that $x = 0$, because $x = -y - z$, because remember that $x, y, z \in W$. So, the statement is $f(x, y, z) = 0$ for $x, y, z \in W$. $x = -y - z - 0$. So, in other words, x, y, z is 0. By 0, I mean the 0 vector in W. So, we have checked it is an isomorphism. So, I should put a tick mark there.

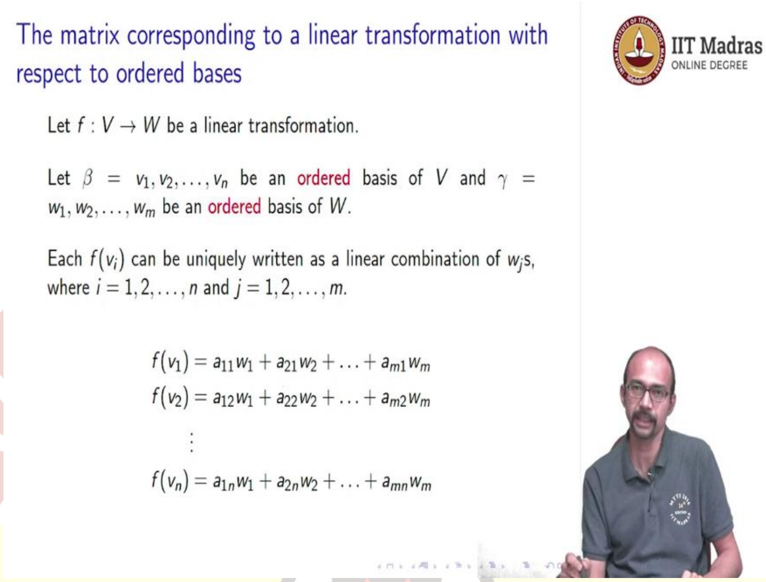
(Refer Slide Time: 11:39)



Let us see how to get the matrix form for a linear transformation. So, first, let us do this example. So, we have $f(x, y) = (2x, y)$. And it is clear that this is the same if we fudge between rows and columns again by writing $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. So, you get $2x, y$, which we write as a row vector as $[2x \quad y]$.

What is the significance of the coefficients in this matrix? So, if we consider the standard basis, 1, 0 and 0, 1 for $\mathbb{R}^2$ and evaluate the function in terms of the basis, we get $f(1, 0) = (2, 0)$, but 2, 0 can be written as 2 times, so you write this vector $(2, 0)$ as a unique linear combination of the basis vectors. So, it is $2(1, 0) + 0(0, 1)$. And similarly, $f(0, 1) = (0, 1) = 0(1, 0) + 1(0, 1)$. And it is clear this 2, 0, 0, 1 that we are getting here is being obtained, those are the numbers obtained in this matrix.

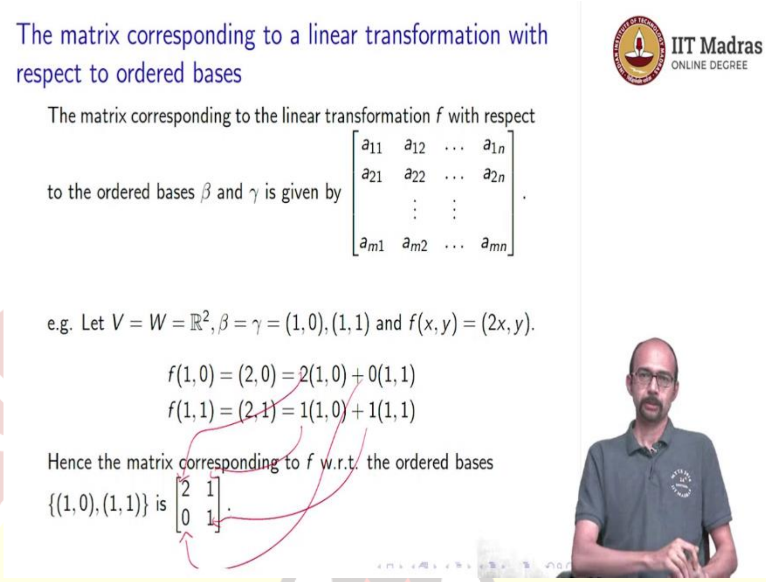
(Refer Slide Time: 13:02)



So, we can make this more general. So, let $f: V \to W$ be a linear transformation. Suppose we have ordered basis for V and W. So, $v_1, v_2, .., v_n$ which we are calling beta is an ordered basis for V. So, what do we mean by order basis? We mean that v_1 is indeed the first vector, v_2 is indeed the second vector. So, I want to distinguish between the basis 1, 0, 0, 1 and 0, 1, 1, 0. This is a standard basis. I mean, both of these can be thought of as a standard, but this is, so this is the ordered standard basis and this is not the ordered standard basis. So, this is the different ordering for the standard basis. So, ordered basis they are different.

So, suppose both of these have ordered this is given to us. And so, we can write f of vi as a unique linear combination with the wjs. So, you write $f(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m$, $f(v_2) = a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m$ and so on, $f(v_n) = a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m$. Then, so notice I am not, I mean, these aijs are written in a particular way. This is more like a_{ji}.

(Refer Slide Time: 14:41)

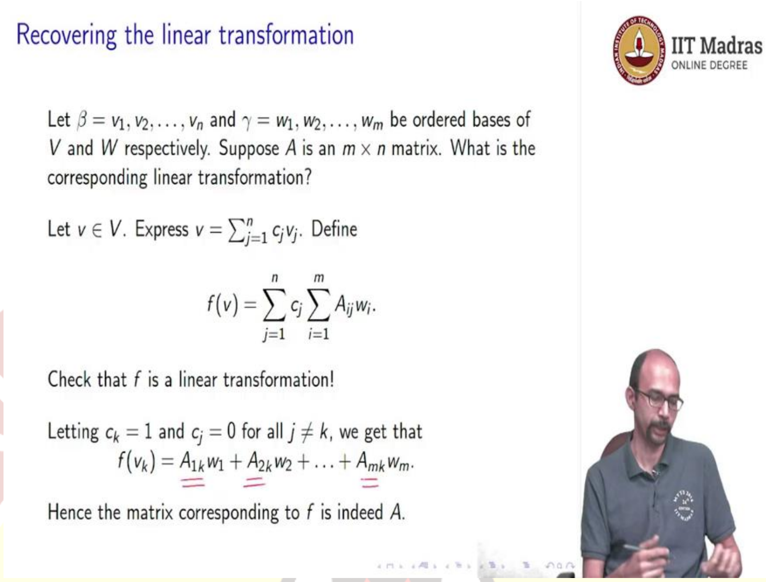


Then the matrix corresponding to this linear transformation with respect to these basis, these ordered basis β and γ which were $v_1, v_2, \dots, v_n$ and $w_1, w_2, \dots, w_m$ respectively is given by $\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$ were the coefficients occurring in the linear combination which express the vector $f(v_1)$. The coefficients in the second, the entries in the second vector, the second column are what occurred in the second linear combination for $f(v_2)$ and so on. The entries in the nth column are what occurred in the expression for the, for $f(v_n)$. So, keep that in mind.

So, just as an example, if you have $V = W = \mathbb{R}^2$ and $\beta = \gamma = (1,0), (1,1)$. So, it is not the ordered standard basis, this one. And let us say $f(x,y) = (2x, y)$. Let us work out what that means. So, you take $f(1,0) = (2,0) = 2(1,0) + 0(1,1)$, and you take $f(1,1) = (2,1) = 1(1,0) + 1(1,1)$.

Then the corresponding matrix is the 2 goes here, the 0 goes here. So, it is in that the first column. And then this 1 goes here, and then this 1 goes here. So, that is what this corresponding matrix is.

(Refer Slide Time: 16:42)



So, can we recover the linear transformation, indeed we can and that is what we are going to study now. Suppose, A is an m by n matrix, what is the corresponding linear transformation. So, let v be in V. We express v in terms of the v_j's. So, express $v = \sum_{j=1}^{n} c_j v_j$. Define $f(v) = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} A_{ij} w_i$. So, here notice that this is slightly different from what we have done before. In our previous thing what we did was, we did, we specify specified values for $f(v_1)$, $f(v_2)$, $f(v_n)$ and then we said you can extend this to a linear transformation.

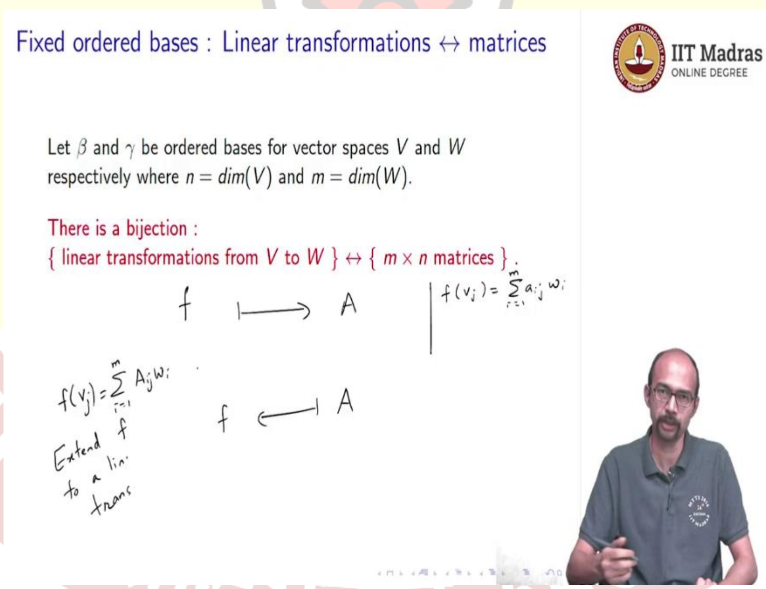
Now, the a priori have not specified what are the values of $v_1, v_2, \ldots, v_n$, meaning what does f evaluated on $v_1, v_2, \ldots, v_n$, instead we have specified ordered basis $w_1, w_2, \ldots, w_m$ and we have specified a matrix A. So, now, what we are going to see is $f(v)$ is whatever expression we have here, $\sum_{j=1}^{n} c_j \sum_{i=1}^{m} A_{ij} w_i$. So, you have to check again this is a linear transformation, not hard. You can write this in terms of in a matrix form and that will maybe simplify things.

But let us evaluate what is $f(v_k)$. So, what is $f(v_k)$? So, $f(v_k)$ can be obtained by putting c_j to be 1 for $j = k$ and $c_j = 0$ when $j \neq k$. And in that case, we get $f(v_k)$ is everything disappears in the sum except c_k which is 1. So, we put $j = k$. So, you get a A_{ik} , $\sum A_{ik} w_i$, which is $A_{1k} w_1 + A_{2k} w_2 + \cdots + A_{mk} w_m$. So, it is this vector which is $f(v_k)$. So, now, we could equivalently have said that we specify the vectors $A_{1k} w_1 + A_{2k} w_2 + \cdots + A_{mk} w_m$similarly for the, as the k^{th} vector. So, we have a set of n vectors like this as k varies and these are the n vectors which are the images of vis. And now we can say the rest is given by extending this f to a linear transformation. That is what we are saying really.

And how do we identify these vectors? Well, this is exactly what the entries in the k^{th} column. So, these are the entries in the kth column. So, the matrix corresponding to f is indeed A, because remember how we get that matrix. To get that matrix, we take the basis $v_1, v_2, \dots, v_n$, we write f of v_i or v_j in terms of the linear combination of w_i's. So, here this is exactly the coefficients and those coefficients are exactly what is going to go into the k^{th} column. So, that is exactly what A is. So, the matrix corresponding to f is indeed A.

So, what have we done here? We saw that for a linear transformation and ordered basis we can get a matrix A that was previously in this side. On this slide, we are saying that suppose we have the ordered basis and matrix A, can we get a linear transformation so that this is the corresponding matrix and indeed we can and this is how we do it.

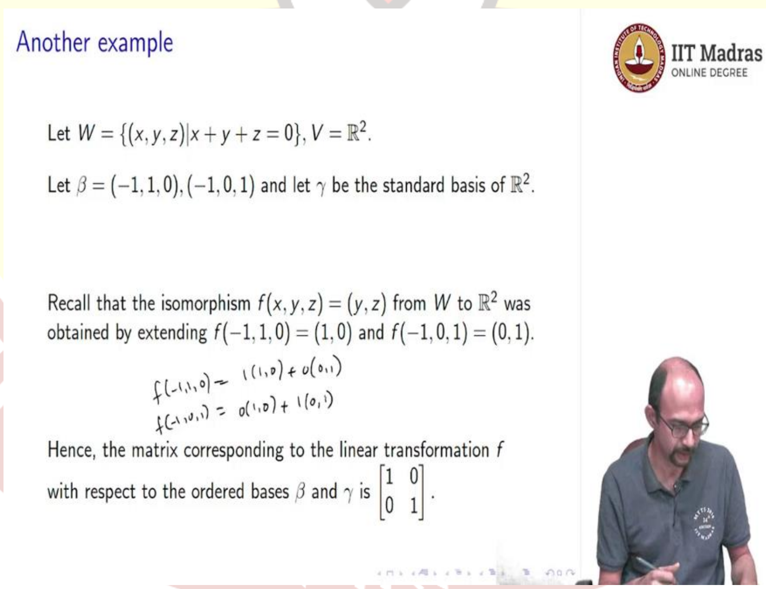
(Refer Slide Time: 20:55)



So, the point is this is actually setting up on a bijection between linear transformations and matrices. So, if beta and gamma are ordered basis for vector spaces V and W and n is the dimension of V and m is the dimension of W there is a bijection linear transformations from V to W to m by n matrices of course, over $\mathbb{R}$. How do we get this? So, if I have a linear transformation, I know how to get a matrix A, because beta and gamma are specified. So, once I know this, I know how to get A. This was by writing $f(v_j) = \sum A_{ij} w_i$. This was, so the A was the matrices, sorry, the matrix A is the ijth entry is this a_{ij}. I have to apologize because I have messed up my notations here. This is $f(v_j)$ is summation over i and not over j. So, j is fixed, s$\sum A_{ij} w_i$. So, this is what these coefficients are exactly what go into the jth column. That is how you get your matrix A.

On the other side, if you have a matrix A, how do I get my f? That is exactly what we saw on the previous slide. By looking at, so we define $f(v_j) = \sum A_{ij} w_i$, this time I should put a capital here, $A_{ij} w_i$. What is this capital A_{ij}? This is exactly the ijth entry of this matrix. And then extend to a linear transformation. That is what we saw earlier that we can do that. So, it is clear that if you start with A, produce f, and then ask what is the corresponding matrix you get back A, that is what we saw on the previous slide. And if you notice, if you see how we have done this, if you start with f and get your matrix and then get the corresponding linear transformation, then you will clearly get back f. So, I will leave that to check to you that this is indeed a bijection. This is the correspondence.

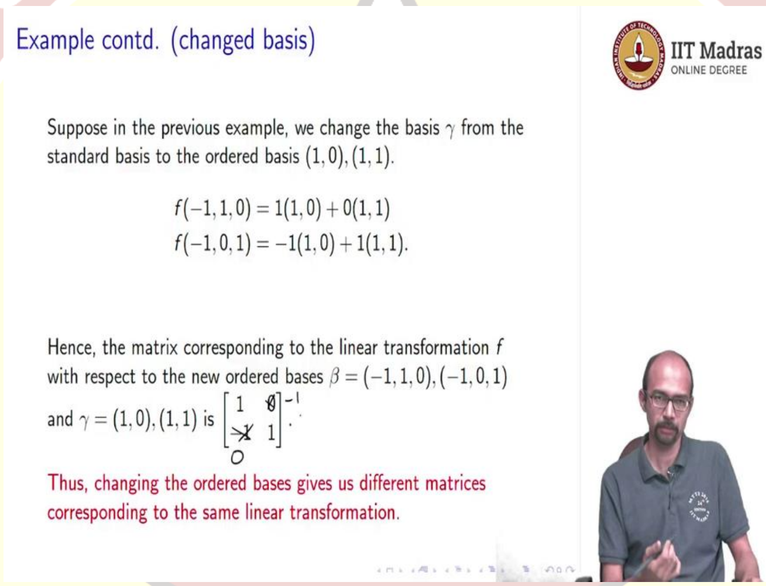
(Refer Slide Time: 23:46)



So, let us do an example not of the correspondence, but of getting the matrix and so on. So, let $W = \{(x, y, z) | x + y + z = 0\}, V = \mathbb{R}^2$. Let β be the basis that we saw, the same ordered basis that we saw before and γ be the standard basis by which we mean the ordered standard basis. So, we had an isomorphism, which was $f(-1,1,0) = (1,0)$ and $f(-1,0,1) = (0,1)$. So, this is a projection to the YZ plane.

So, the matrix corresponding to this linear transformation f with respect to the ordered basis β and γ is just (1, 0), (0, 1), the identity matrix. Why is that? How do we get the matrix? You look at f evaluated on the basis vectors and then write that in terms of the basis vectors for the range, space. So, in this case, $f(-1,1,0) = 1(1,0) + 0(0,1)$, and similarly, $-1,0,-1\ is\ 0(1,0) + 1(0,1)$ and that is how you get your identity matrix.

(Refer Slide Time: 25:03)

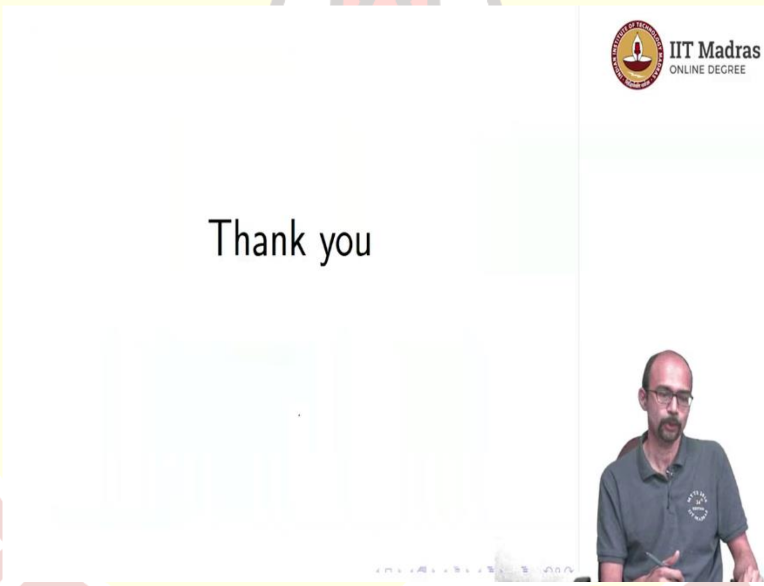


But now suppose we change our basis. So, instead of taking the standard ordered basis, we take the basis (1, 0), and (1, 1, 1) γ. So, we have the, remember, we have the same linear transformation, but we are changing our ordered this. Let us see what happens. So, now if we write $f(-1,1,0)$ in terms of the basis, ordered basis γ, we get, so remember that this is $(1,0)$, that is $1(1,0) + 0(1,1)$. So, no change in the coefficients that were coming in the first equation earlier. But for the second one, we have $f(-1,0,1) = -1(1,0) + 1(1,1)$.

So, what is the corresponding matrix? The corresponding matrix is you put these coefficients in the first equation as the first column, so this is not correct what I have written here, in the first column, so this is 1 and then this is wrong and what should have been there is 0 and then again, here, this should have been the -1 and the 1. But the main point here is in the previous slide we saw that we got the identity matrix with respect to the ordered basis beta and a different basis γ, the standard ordered basis, whereas here we get a different matrix when we change the basis.

So, your basis, what basis you choose is very important. So, it will not happen that you will change basis and get the same matrix. With different ordered basis you will get different matrices. But if you keep your basis constant, then there is a bijection between linear transformations and matrices. So, you will get the same matrix in particular. So, the important part here is that you will get, is that changing ordered basis gives you different matrices corresponding to the same linear transformation. So, it is a good question to ask, if you did a linear transformation what is a possible set of matrices that you can get and this is something interesting and we will be looking into that in a video later on.

(Refer Slide Time: 27:36)



Thank you.