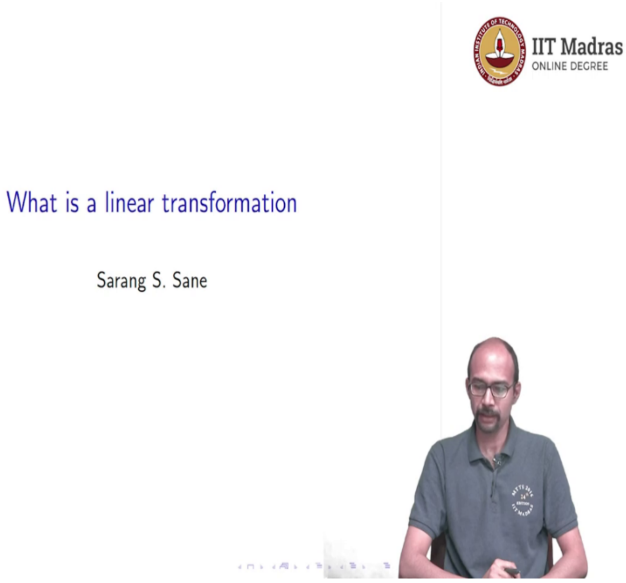
IIT Madras

ONLINE DEGREE

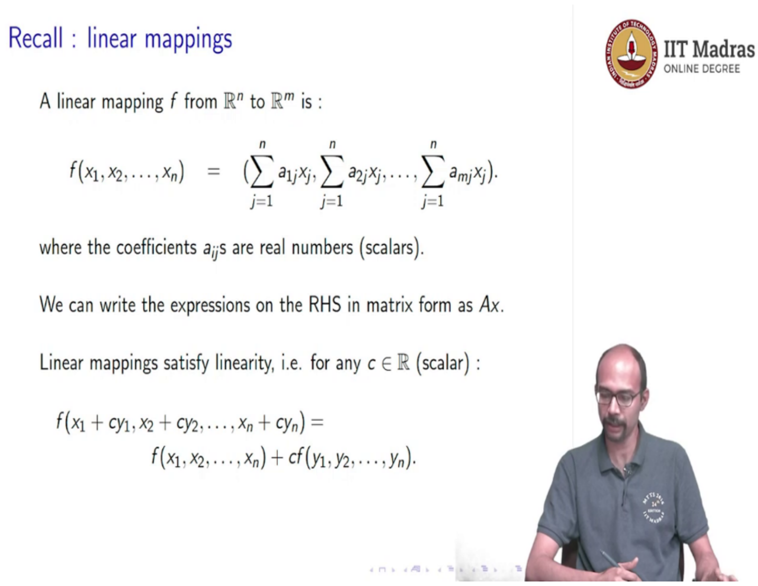
Mathematics of Data Science 2
Professor Sarang Sane
Department of Mathematics,
Indian Institute of Technology, Madras
What is a linear transformation

(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 study what is a linear transformation.

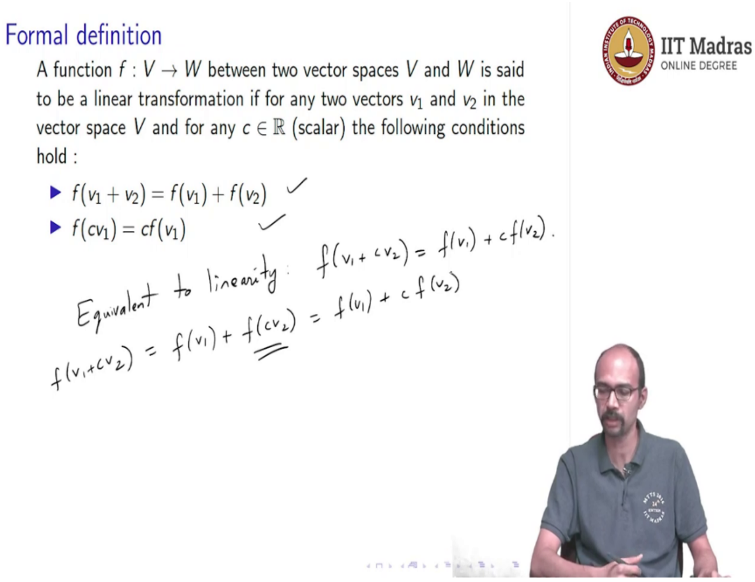
(Refer Slide Time: 00:27)



So, let us just recall before we go ahead, what is a linear mapping, this is what we studied in the previous video. A linear mapping f from $\mathbb{R}^n$ to $\mathbb{R}^m$, so it is a function, and the function has a very particular form. Namely, the form is that every coordinate is given by a linear combination. This is the form of the function, and linear combination with what coefficients, the coefficients are real numbers.

So, we studied this to some extent in the previous video, and we saw the use of such a thing. For example, we looked at the examples of the grocery shops. And we noticed that linear mappings have some very interesting properties called linearity. So further, we can write this in the form of matrix multiplication, Ax , where the $(ij)^{th}$ entry of the matrix A is a_{ij} and x is the column vector with $x_1, x_2, .. x_n$ and linear mapping satisfy linearity, which means $f(x_1+cy_1, x_2+cy_2, .., x_n+cy_n)=f(x_1, x_2, .., x_n)+cf(y_1+y_2+..+y_n)$.

(Refer Slide Time: 01:55)

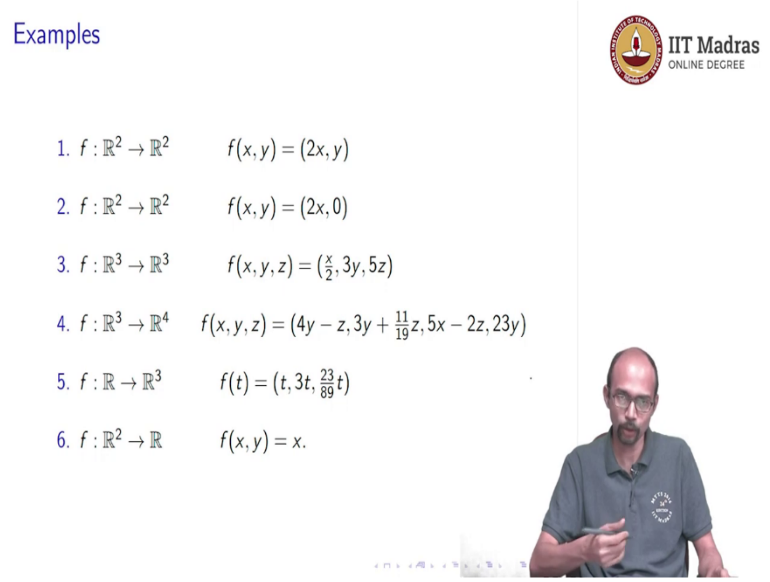


So, let us define what is a linear transformation. So, a linear transformation is basically whatever we want the same properties that we had in the previous slide, but for arbitrary vector spaces, instead of for $\mathbb{R}^n$ and $\mathbb{R}^m$ we look at vector spaces V and W. So, we say, a function f from V to W is a linear transformation, if for any two vectors, v_1 and v_2 in the vector space V, and for any $c \in \mathbb{R}$ so that is a scalar the following conditions hold: $f(v_1+v_2)=f(v_1)+f(v_2), f(cv_1)=cf(v_1)$.

So before going ahead, maybe, I will point out that this is the same as linearity. So, this is equivalent to linearity as in the previous slide. Why is that? Suppose we assume it has linearity, then, of course, both of these are clear, because we can put $c=1$ and that will give us the first condition $f(v_1+cv_2)=f(v_1)+cf(v_2)$. By taking $v_1=0$ we get the second one. So, if we know linearity is true, then we get both these that both these conditions hold. Conversely, if we know both of the conditions hold, so I do not know linearity a priory, but I know that these two conditions hold.

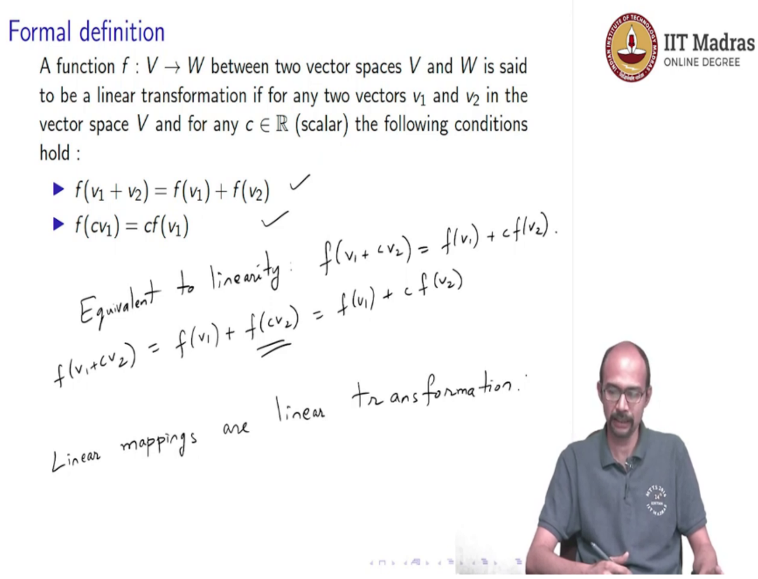
So, I can treat cv_2 as a vector, and use the first condition to write this as as $f(v_1)+cf(v_2)$. We use the first equation with cv_2 in place of this v_2. And then we use the second equation, the second condition for writing the second term here, $f(cv_2)=cf(v_2)$, and that is what we want for linearity. So, this is equivalent to linearity. So, you could either express it this way or you could express it in terms of the way we have written it, which is we have written it earlier.

(Refer Slide Time: 04:36)



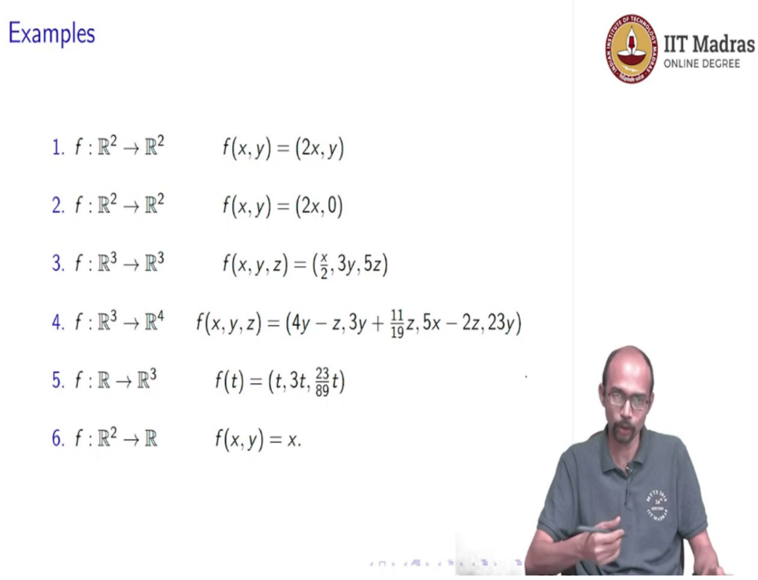
Let us look at some examples. So, here are some examples. So, these are all going to be examples of linear mapping. So, in particular, they are linear transformations. So, I should have pointed out before that linear mappings are linear transformations because linear mapping satisfy linearity.

(Refer Slide Time: 04:59)



So maybe I should write here. So, linear mappings are linear transformations as we have defined above. So, linear mappings mean when we have V and W, as some $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and then the function f is given by a bunch of a collection of linear combinations. We have seen already that in that case that satisfy linearity, and we have checked linearity is the same as these two conditions.

(Refer Slide Time: 05:32)

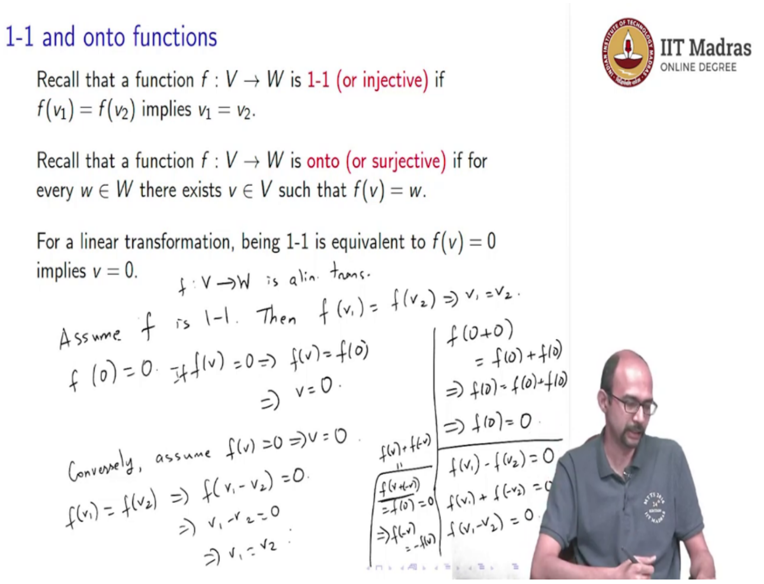


So, all these examples shown in this slide are examples of linear mapping. So, in particular, they are linear transformations.

And maybe a last function which is also linear transformation, which is a projection function $f(x,y)=x$, so it projects onto x. We will see this later in some other context a few weeks from now.

So, all these are linear mapping so in particular they are linear transformations. Of course, we will also see some non-trivial examples of linear transformations. By non-trivial I mean, where the V or W are not $\mathbb{R}^m$ or $\mathbb{R}^n$

(Refer Slide Time: 06:43)



So, let us recall before going ahead what is a one-one function and what is an onto function. Possibly you have seen this already Maths 1, if not, just look at this definition. So, function $f:V\to W$. Now V and W, need not be vector spaces here these could be any sets. So, it is called one-one or injective if $f(v_1)=f(v_2)\Rightarrow v_1=v_2$.

Similarly, such a function is called onto or surjective if for any $w\in W, v\in V$, such that $f(v)=w$. So, for a linear transformation being one-one is equivalent to f of V0 implies V0. Why is that? So, let us prove that.

So, let us assume that f is one-one. Notice f is a linear transformation. I am assuming it is a linear transformation. Otherwise, of course, none of this even makes sense because we do not even know what is the 0. So, one of the properties of a linear transformation is that $f(0)=0$. So here what do we mean by 0 and 0 here, this, the 0 inside the bracket is the 0 in V it is a vector space, remember, so it has a value of 0 in W. So, this is true because it is a linear transformation.

Maybe, I should point out why for a linear transformation $f(0)=0$ that is because if you take $f(0+0)=f(0)+f(0)$. So, we can, so this is happening inside a vector space W, so we can subtract $f(0)$ from both sides.

So, what that will give us is that $f(0)=0$. So, we know that, so this is true for any linear transformation. So now we want to check that $f(v)=0 \Rightarrow v=0$. So, that means $f(v)=f(0)$. And now I can use the fact that f is one-one to conclude that v is 0. So, this shows that if f is one-one, then this condition is satisfied that $f(v)=0 \Rightarrow v=0$.

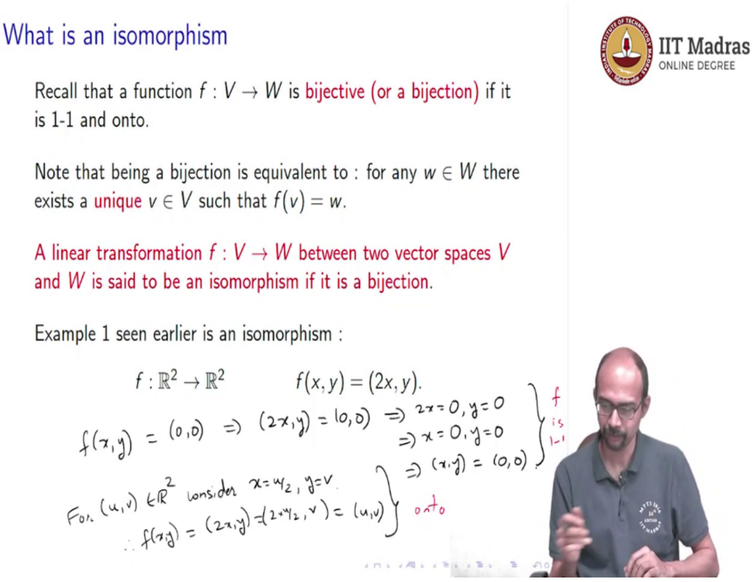
Conversely, we assume that assume $f(v)=0 \Rightarrow v=0$. So now I want to prove it is one-one. So, let us assume that $f(v_1)=f(v_2)$, but then that will mean $f(v_1-v_2)=0$.

So, again, here I have to explain why. So that is because I can subtract $f(v_2)$ on both sides so I will get $f(v_1)-f(v_2)=0$.

And then I am using a linearity because it is a linear transformation to say that $f(v_1-v_2)=0$. So, the key step is to is prove that I can take minus inside. Why can I take minus inside? So, the reason is, the same argument that we had before. So, if we have $v+(-v)$, then this is $f(0)=0$, which we already proved.

So, in general, if you have a arbitrary function on sets, checking one-one means you have to check this condition that $f(v_1)=f(v_2) \Rightarrow v_1=v_2$. But if you have a linear transformation, you can instead just check that $f(v)=0 \Rightarrow v=0$. So, it reduces to checking for those values of v, so that when you apply f to those values, they become 0. So, this is a special set and we are going to study that set later. In fact, it is a subspace.

(Refer Slide Time: 13:19)



So now let us get to what is an isomorphism, which is why we studied this notion of one-one and onto. So, recall first that a function f, again, for any arbitrary sets, V and W, if you have a function f from V to W it is called a bijection or we say that it is bijective; if it is one-one and onto. So, or sometimes it is called a one-one correspondence. What that means is, if you have v here, then there is a unique w that corresponds to it, which is what we are calling $f(v)$.

So, there is exactly one such v, such that $f(v)=w$. For each w such a v exists, and it is unique, that is what is called a bijection or that is the essence of a bijection. So in other words, you can pair up elements, you have v here, w here, v here, w here, that those pairs are exactly, the w paired with v is what we are calling $f(v)$.

So that is, if you can do such a thing that is called a bijection. Fine, so what is an isomorphism? So, this is what I just said, bijection is equivalent to saying that for any $w \in W$, $v \in V$ there exists a unique v such that . So, a linear transformation between two vector spaces V and W is an isomorphism if it is a bijection. So, it is an isomorphism if it is a bijection. So f satisfies linearity, which means $f(v_1+v_2)=f(v_1)+f(v_2), f(cv_1)=cf(v_1)$ and it satisfies that it is one-one and onto.

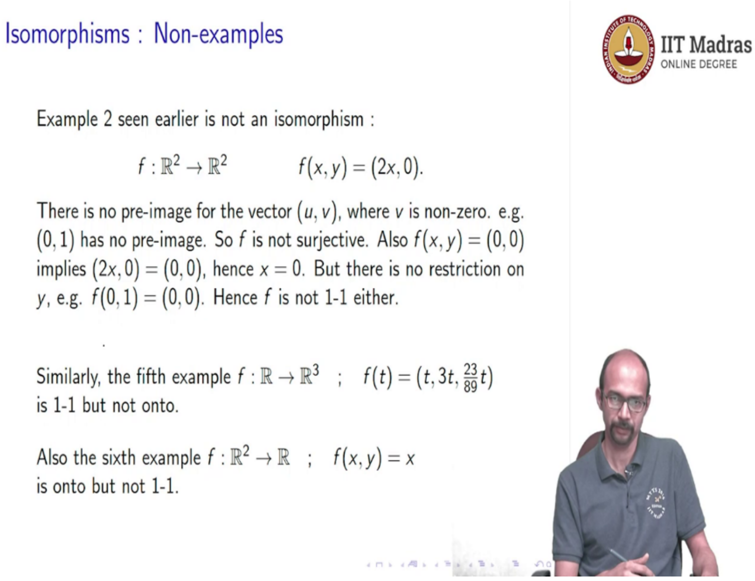
Namely, that for every $w \in W$, there is a unique little $v \in V$, such that $f(v)=w$. So, example one seen earlier is an isomorphism, let us work that out. So, what was the example, it was $f:\mathbb{R}^2 \to \mathbb{R}^2 f(x,y)=(2x,y)$, we already know this is a linear transformation because it is a linear mapping.

So, to see that it is a bijection, so because it is a linear transformation to check injectivity it is enough to check that whenever this becomes 0 and remember that 0 in $\mathbb{R}^2$ is the vector (0, 0). That means, $(2x, y)=(0,0)$.

But when will this be true? This will be $(0,0)$ precisely when our $2x=0, y=0$. So, I can see that, which means $(x,y)=(0,0)$. So this is saying $f(x,y)=0 \Rightarrow (x,y)=(0,0)$ So, we have checked it's injective, or one-one.

Let us check that it is surjective or onto. To check that it is onto we have to show that if you are given any element let us say $(u,v)\in\mathbb{R}^2$, you can get (x,y) such that $f(x,y)=(u,v)$. So, if so, consider $x=u/2, y=0$. So, this is now saying it is onto. Great. So, we have seen how to prove that this is an isomorphism.

(Refer Slide Time: 17:47)



So, we should also look at examples of things which are not linear transformations which are not isomorphisms. So, here is an example, which is not an isomorphism. The second example that we saw. So, this is $f(x,y)=(2x,0)$. So, what goes wrong? Well, several things go wrong. First of all, there is no pre-image for the vector (u,v) where v is non-zero.

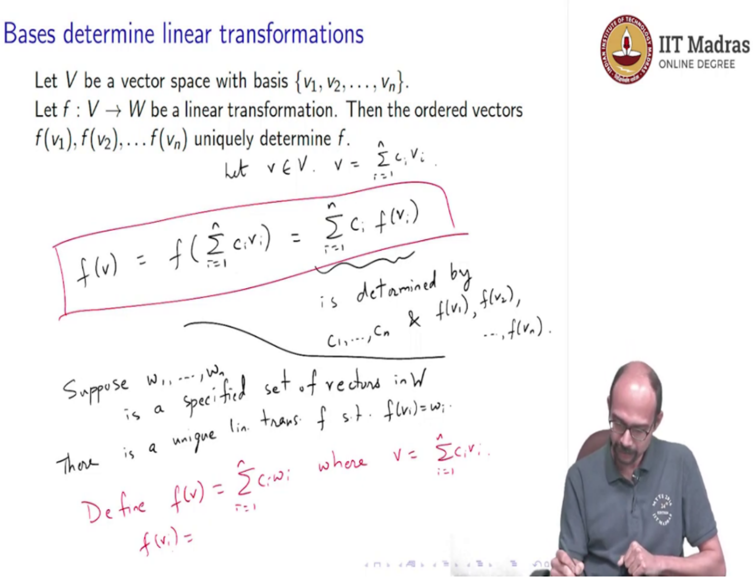
For example, (0,1) has no pre-image. Why does not it have a pre-image because you can see that the only vectors that are images are of the form something comma 0, so (0,1) cannot have a pre-image. So that means there is no $(x,y)\in\mathbb{R}^2$ such that $f(x,y)=(0,1)$ that is what we are saying. So, this is not onto or surjective.

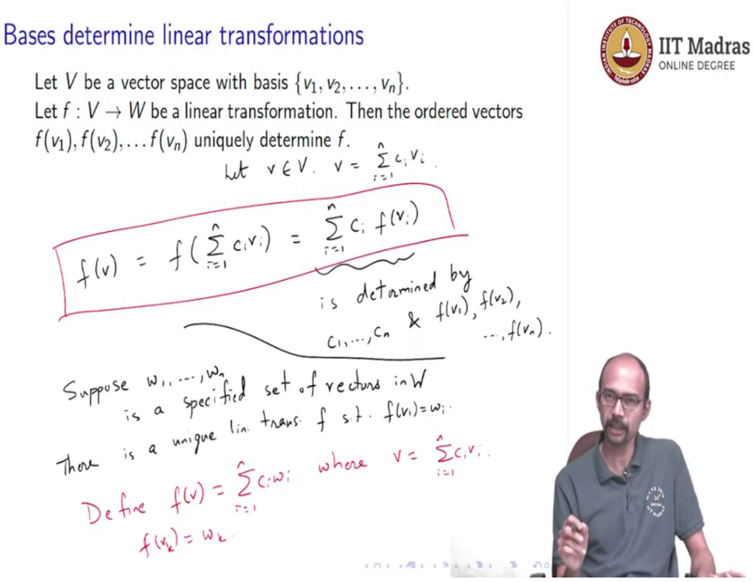
So similarly, if you take f of x, y to be 0, 0 then that means to 2x is 0, 2x comma 0 is 0 comma 0, which means x is 0, but y could be anything. So, for example, if you take f of 0 comma 1, then the answer you get is 0. That means, we cannot say that f of x, y is 0, 0 implies x, y is 0, 0 because you could have many other choices.

For example, you could have $(0,1)$. So, this is not one-one. So, both of the conditions needed for a bijection fail. So, this is a, this is not an isomorphism by any margin. There can be examples where you have one-one but not onto or onto but not one-one. Here is an example which is not onto. So, $f(t)=(t,3t,\frac{23}{89}t)$, this is one-one but not onto, I will suggest you check why it is so.

And similarly, if you take the projection map $\mathbb{R}^2 \to \mathbb{R}$, then this is onto but not one-one. Again I will suggest you check why it is so. So, I hope, this gives you a feeling for what we mean by an isomorphism.

(Refer Slide Time: 20:09)





So, let V be a vector space with basis $\{v_1, v_2, .., v_n\}$. So, now, let us introduce basis into the picture. So, the what we are going to say here is that if you have a basis then the linear transformation is determined by the values it takes on the basis. So, suppose $\{v_1, v_2, .., v_n\}$ is a basis and f is a linear transformation, then the ordered vectors $f(v_1), f(v_2), .., f(v_n)$ uniquely determine f .

Why do I say this? So, the reason one says this is the following. If I want to get $f(v)$. So, I am claiming that it is determined by $f(v_2), f(v_3), \ldots, f(v_n)$. So, we know that if $\{v_1, v_2, \ldots, v_n\}$ is a basis, there is a unique linear combination corresponding to each which equals any given vector in that vector space. So $v = \sum_{i=1}^{n} c_i v_i$ for a unique set of scalars $\{c_1, c_2, \ldots, c_n\}$.

So, then what is $f(v)$? It is $f(\sum_{i=1}^{n} c_{i0} v_i)$, but now, I know that f is a linear transformation. So, in particular I mean that means it satisfies linearity. So, by linearity, I can pull out the summation and the scalars and write this as $\sum_{i=1}^{n} c_i f(v_i)$. I do this sequentially. So, in my definition of linearity, I had $f(v_1 + cv_2) = f(v_1) + cf(v_2)$. So, I keep applying it to. First I applied to v_n and then I applied to v_{n-1} and so on. And I can so, I can pull out any sum. It need not be only vectors it could be any finite sum. And similarly, I can pull out scalars that is what the second condition of linearity told us.

So, since that is the case, this $f(v)$ is determined by the values of $f(v_2), f(v_3), \ldots, f(v_n)$. By values I mean the vectors, these are vectors, not numbers. So, if you take a vector v, and I want to evaluate $f(v)$, all I need to know are what are the scalars c_i and what are the vectors $f(v_2), f(v_3), \ldots, f(v_n)$, that determines $f(v)$, and that is what we mean by saying that a basis determines an ordered basis mind determines the linear transformation.

I mean, we could flip this around, and say that suppose I specify $f(v_1), f(v_2), \ldots, f(v_n)$ what are I mean, some n many vectors . So, suppose $w_1, w_2, \ldots, w_n$ is a specified set of vectors in W then there is a unique linear transformation if such that $f(v_i) = w_i$. How do I define this linear transformation? That is exactly what this equation here is telling.

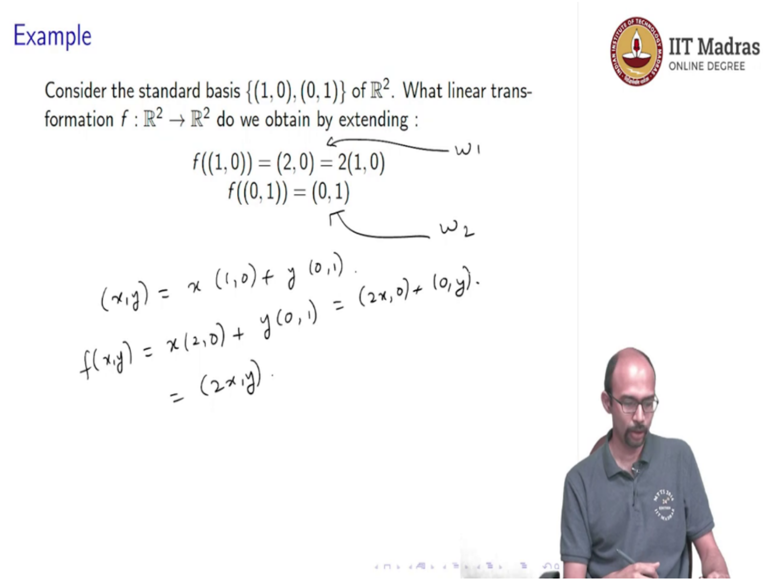
So, what I do is I define f to be so define $f(\sum_{i=1}^{n} c_i v_i) = \sum_{i=1}^{n} w_i$. And then let us see what is $f(v_i)$. So, to get any $f(v_i)$, you have to first see how is v_i expressed in terms, how is that vector expressed in terms of the basis, but in terms of the basis v_i itself is a basis vector.

So, maybe let us say v_k .So, you put $c_k = 1$, and all the other $c_i = 0$, and that is the only way in which you can write, v_k as a linear combination of these basis vectors, because it is a basis.

So, if you do that, then all the terms on the in this expression, submission, $c_i w_i$ will vanish except the kth one. And for the kth one, what you will get is $c_k = 1$, so you will get exactly w_k.

So, $f(v_k) = w_k$, that is what we wanted. So now this linear, so this linear transformation satisfies indeed that $f(v_i) = w_i$. And it is unique because every linear transformation is determined by what values the basis vectors take. And since they take I mean, these values are specified, no other linear transformation can force that $f(v_i) = w_i$, that is exactly what we proved before. So with this, let us go ahead. This is a slightly tricky point and I mean, I want a bit you to think a little about this as we go ahead.

(Refer Slide Time: 26:21)

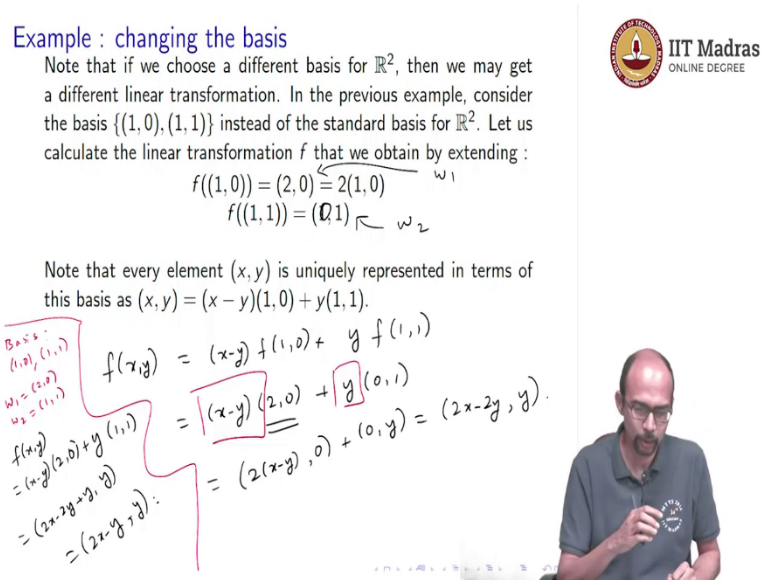


Let us do this on a particular example. Suppose I take the standard basis $\{(1,0),(0,1)\}$ for $\mathbb{R}^2$.

So here, in this example, we have, these two vectors (2, 0) and (0,1) and let me write that so this is your w_1 and this is w_2. So we want to extend this function f to a linear transformation from $\mathbb{R}^2 \to \mathbb{R}^2$.

So, to do that, we have to write any (x,y) first as a linear combination of (1,0) and (0,1). $x(1,0)+y(0,1)$. And then I define $f(x,y)=f(x(1,0))+f(y(0,1))=x(2,0)+y(0,1)$. So, this is (2x,y). This was exactly the first example that we looked at, in this video of linear transformations or linear mappings. So, this is the function that we get by or the linear transformation we get by extending the function, which takes (1, 0) to (2, 0) and (0, 1) to (0, 1) that is the way we say this.

(Refer Slide Time: 28:21)



So, let us change the basis and see what we get if we change the basis. So, the point here is if you change the basis, you will get a different linear transformations, and this is very important. So, in the previous example, instead of working with the standard basis let us work with the basis $\{(1,0),(1,1)\}$. So let us say take (1, 0) to (2, 0) and (1, 1) to (0, 1).

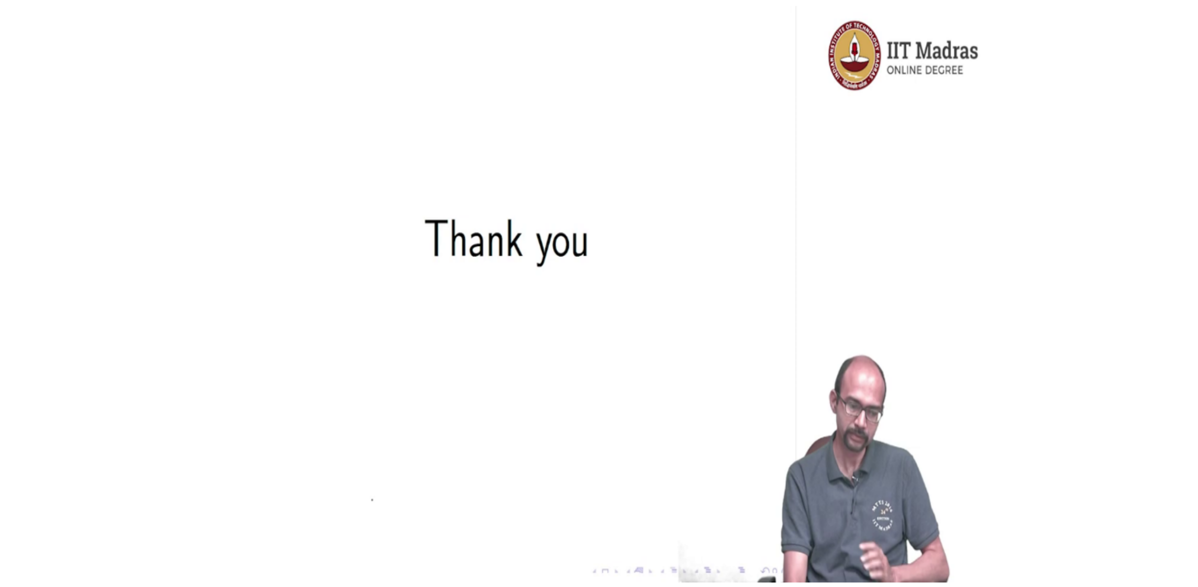
So, if that if you do that, then what do we get? So, every element (x,y) is uniquely represented in terms of this basis as $(x-y)(1,0)+y(1,1)$. So basically the basis was changed, but the values were the same. So here we have w_1,w_2 , they are the same, but the basis was changed. And the question is what linear transformation do you get by extending? So let us apply the same thing.

So, we have $f(x,y)=(x-y)(1,0)+y(1,1)$. So, this is $(x-y)(2,0)+y(0,1)=(2x-2y,y)$. So, note that w_1, w_2 are the same. So, this part will, meaning these two things will remain the same. But what changes?

What changes are the coefficients . So, y has remained the same, but x has changed.

If instead we had chosen, I mean, just as an example, since I had that typo over there, let me work that out as well. If instead, my basis was the one that I have here, (1, 0) and (1,1) and $w_1=(2,0), w_2=(1,1)$, then what would have (x, y), have been? So $f(x,y)=(x-y)(2,0)+y(1,1)=(2x-y,y)$, if you work out. So this is a second example.

(Refer Slide Time: 31:38)



So, let us just summarize what we have done this video. So, in this video, we have seen what is the linear transformation. Essentially, it is just generalizing the definition of a linear mapping, which was a bunch of linear combinations to arbitrary vector spaces. Of course, in arbitrary vector spaces, we do not have the, I mean linear combinations, we have to make sense of what that means.

So, instead of doing it that way, we specified by the other property, which is to say that they are satisfying the two conditions of linearity or equivalently a single condition that we call linearity in the previous video. So that is what was a linear combination, we saw, of course, just by its very nature, that linear mappings are linear transformations. And we saw a bunch of examples.

And we saw that they can be written in some kind of matrix form. And finally, we saw the relevance of a basis to all this. Namely, that you can use the basis to express a linear transformation in terms of the values that are taken by the basis vectors, and that actually determines the linear transformation entirely.

And finally, we saw that if you, you have to be careful about what basis you choose. And so, if you choose different basis the linear transformations you are going to get by extending the basis may be different, so that you have to be careful about. So, thank you.