INDIAN INSTITUTE OF TECHNOLOGY MADRAS

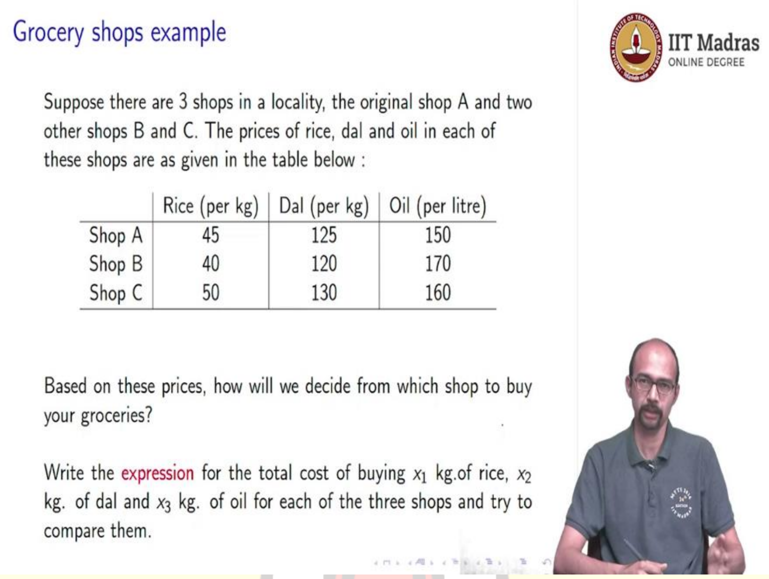
सिद्धिर्भवति कर्मजा

IIT Madras

ONLINE DEGREE

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

(Refer Slide Time: 00:14)

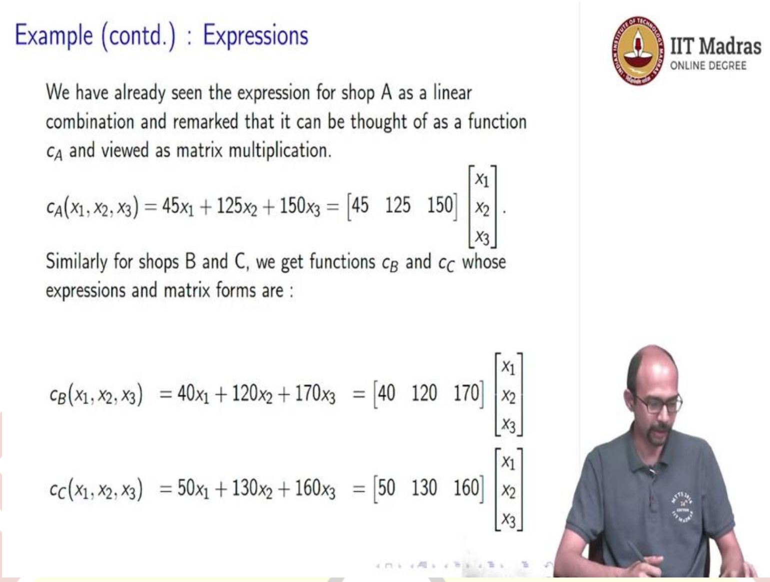


Let us move on to a slightly more elaborate example. So, we will use the same data that we have before, we have shop A. But now in this locality Malgudi, on this town Malgudi, we have the original shop A, and we have two other shops B and C as well, and the prices of rice, dal and oil in each of the shops is given in this table.

So, in Shop A we already have seen that it is 45 kgs, 45 rupees per kg for rice, 125 rupees per kg for a dal and I think 150 rupees per liter for oil. And for Shop B we have 40, 120 and 170 respectively and Shop C we have 50, 130 and 160 respectively.

Now, someone asked the caterer why do you always buy from shop A? So, is there a good way of deciding, which shop to buy your groceries from? So, in order to do this, what do we have to do, we have to check out the cost for buying x_1 kgs of rice, x_2 kgs of dal and x_3 liters of oil from each of the shops, then compare them and see which one is less, and then we would go there. Of course, let us assume that all of them have the same quality and there is no difference as such.

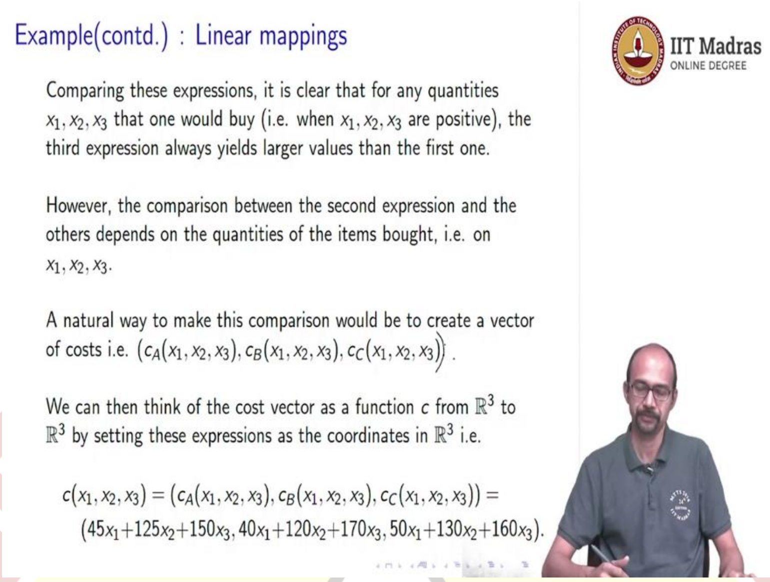
(Refer Slide Time: 01:44)

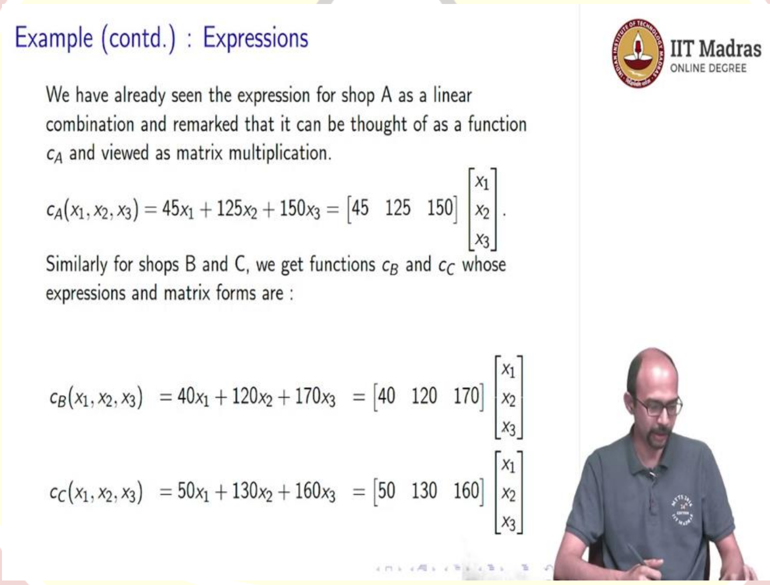


So, let us do that. So, we have seen already that the expression for shop A is a linear combination, and we can, in fact view it as a function. That is what we saw in a few slides ago, and we can also think of it as matrix multiplication. Just to remind you, we had this thing called c_A, which was the function, which was this linear combination $45x_1 + 125x_2 + 150x_3$, and which we can also write as a matrix multiplication.

So similarly, for shops B and C, we can write the cost function c_B and c_C. So, the shop is capital C and the cost is little c. And expressions and matrix forms are as below. So, c_B is the linear combination $40x_1 + 120x_2 + 170x_3$ and that is the corresponding matrix multiplication. And for shop C, we have the function c_C, which is $50x_1 + 130x_2 + 160x_3$ and the corresponding matrix multiplication.

(Refer Slide Time: 02:57)

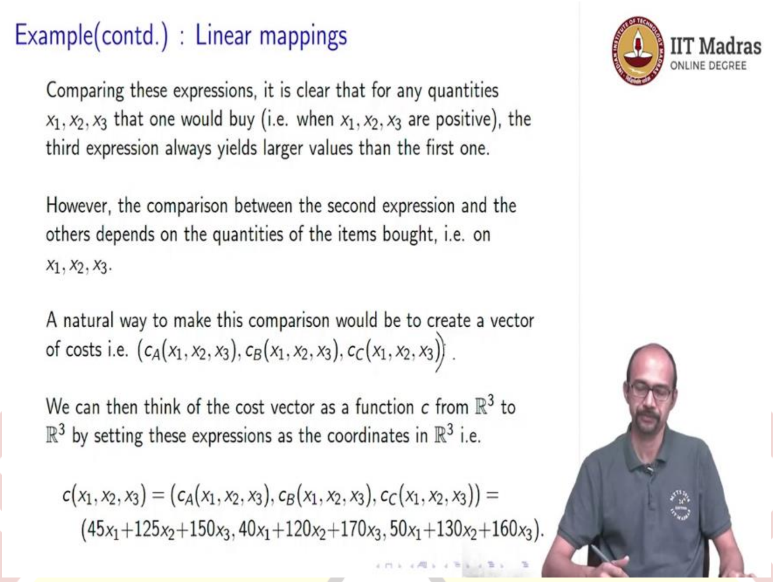




So, to compare these expressions, let us go back for a second and check them out. We had $45x_1 + 125x_2 + 150x_3,\ 40x_1 + 120x_2 + 170x_3$. So, rice and dal are cheaper in B than in A, but oil is more expensive. And then we have $50x_1 + 130x_2 + 160x_3$. So, rice and dal in C are more expensive than in shop B, but oil is less expensive.

However, when you compare it with shop A, all three are more expensive for shop C. So, you would, assuming everything else is the same you would always prefer shop C, shop A over shop C. So that is what we see here.

(Refer Slide Time: 03:51)



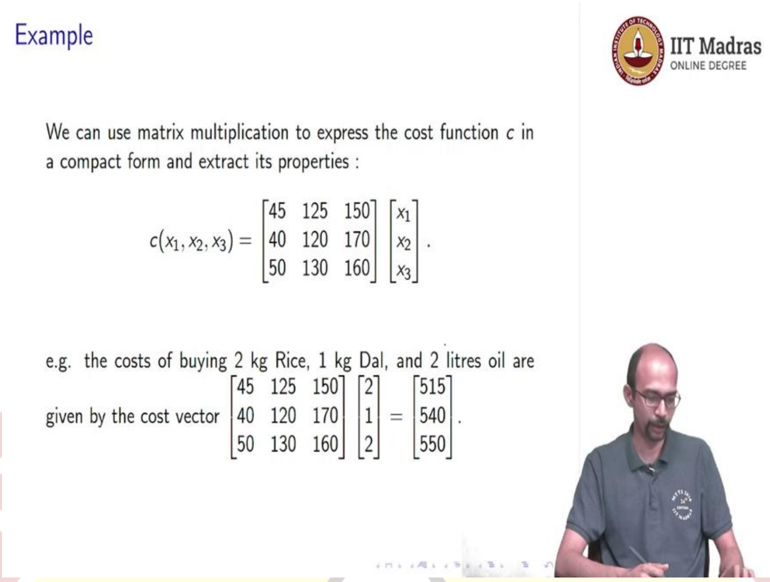
The third expression always yields larger values than the first one. Of course, we are talking about commodities, so we think of them as buying the, I mean, buying commodities from a grocery shop. So x_1, x_2, x_3 are always I think of them as positive numbers, but we do not know how to compare between the second shop that is shop B and shops A and C.

So, to compare this, we would want to compare these cost functions. So, the idea you would I mean, the natural way of doing this would be to treat them together c_A, c_B and c_C we could think of them as a vector. So, you look at this vector, c_A, c_B, c_C and then we can think of this actually as a function. Because each of these, after all are functions from $\mathbb{R}^3$ to $\mathbb{R}$, so this vector can be thought of as an element of $\mathbb{R}^3$, it is a point of $\mathbb{R}^3$.

So, we can think of the cost function which is $c(x_1, x_2, x_3) = (c_A(x_1, x_2, x_3), c_B(x_1, x_2, x_3), c_C(x_1, x_2, x_3)) = (45x_1 + 125x_2 + 150x_3, 40x_1 + 120x_2 + 170x_3, 50x_1 + 130x_2 + 160x_3)$.

So, this is the function, the cost function, wherein the first coordinate of the cost function tells you the cost in shop A, the second coordinate tells you in shop B, and the third coordinate tells you in shop C. And now, when you are given x_1, x_2, x_3 you put it into this cost function, look at these three numbers that it produces and then you choose the smallest one, you can just read them off. So, this would be a good way of trying to compare. Of course, there are many other reasons to look at the cost function.

(Refer Slide Time: 06:10)

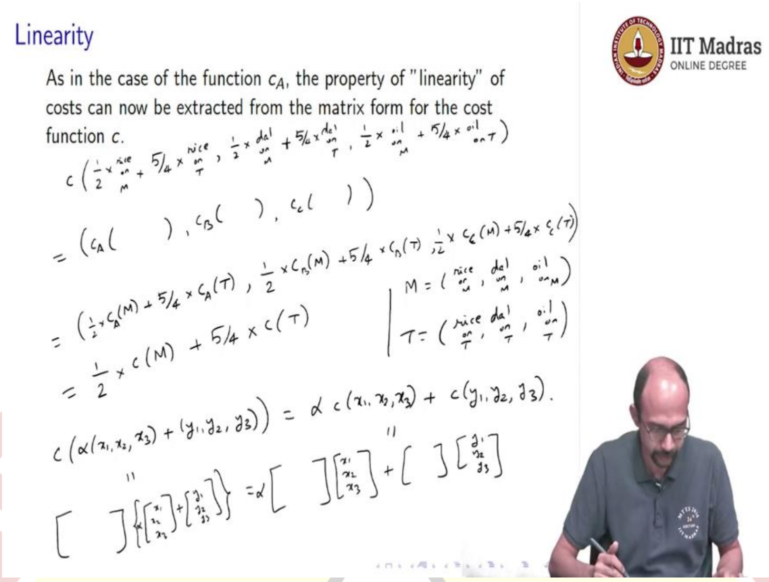


For example, we can treat this as matrix multiplication. Because after all each of these are given by matrix multiplication, where we have a column vector x_1, x_2, x_3 on the right and a row vector corresponding to the cost of individual items on the left. So, you can put these three row vectors together and get this matrix.

So, $c(x_1, x_2, x_3)$ has a matrix form where the matrix is 45, 125, 150. This was the prices for shop A, 40, 120, 170, which are the prices for shop B and for 50, 130, 160, which are the costs for shop C. So, this is exactly what your table was the cost table, a few slides ago. And you can see by just multiplying out that this is indeed what the cost vector is.

So, just as an example, the cost of buying 2 kgs of rice, 1 kg of dal and 2 liters of oil are given by the cost vector, you put $x_1 = 2$, $x_2 = 1$ and $x_3 = 2$ and we will get 515, 540 and 550, if hopefully I have done this computation correctly.

(Refer Slide Time: 07:30)



So as in the case of the function c_A, the property of linearity of costs can now be extracted from the matrix form for the cost function c. So, let us first of all understand this statement. What do we mean by the statement, the statement means that if you buy, so suppose again, we had a situation like Monday and Tuesday, and then we had a say we wanted to know what the costs for each shop on Wednesday are, so I would be evaluating this on half times the rice on Monday, plus 5 by 4 times rice on Tuesday comma halftimes rice, dal excuse me, dal on Monday plus 5 by 4 times dal on Tuesday, and then half times oil on Monday plus 5 by 4 times oil on Tuesday.

You would be evaluating it on this, which means you have this matrix. Well, let me come to the matrix in a second, which means basically you are doing c_A of the same thing, c_B of the same thing and c_C of the same thing, but we saw that c_A is basically after we finished the computations the c_A of this is exactly half times c_A on Monday. So, I will just call it Monday.

By Monday, what do I mean? I mean rice on Monday, the vector rice on Monday, dal on Monday comma oil on Monday, plus 5 by 4 times c_A on Tuesday, which means by T I mean the vector rice on Tuesday, dal on Tuesday, oil on Tuesday. And then I can, the same way as I did for c_A the same. There is nothing special about the shop A. So, the same thing works for c_B as well and for c_C as well.

So, I have this very nice expression, and then half times c_C of Monday plus 5 by 4 times c_C on Tuesday. And so, I can effectively compute this by computing everything on Monday and

Tuesday, but it gets even better because we know how to do matrix. We know how to do, how to add and subtract and scalar multiply things in $\mathbb{R}^3$.

So, because we can do that, we can write this in even better as half times the cost vector evaluated on the commodities for Monday, plus 5 by 4 times the cost vector evaluated on the vector for Tuesday. I can separate these into two different vectors, and then I can pull out the scalars, I will encourage you to do that. I will just, just in case it was confusing what is M. By M I meant rice on Monday then dal on Monday, oil on Monday. And by T I meant rice on Tuesday, dal on Tuesday, oil on Tuesday.

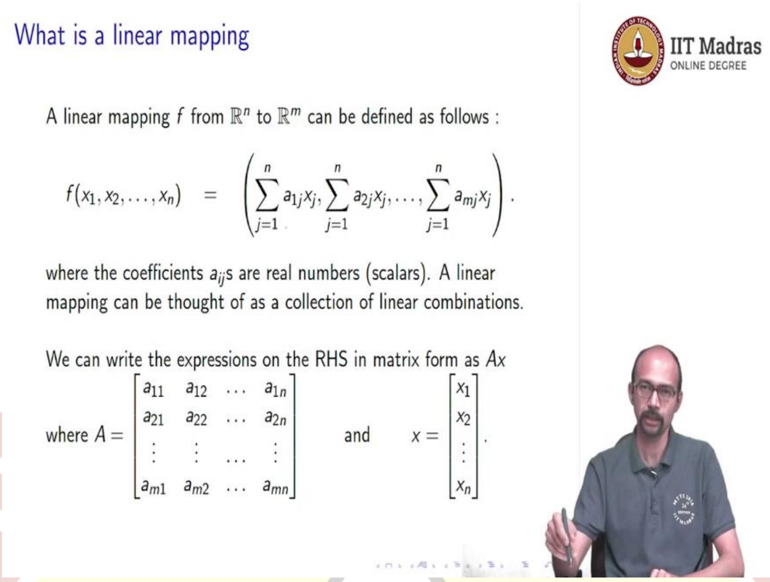
So, notice, I do not even remember what the quantity is there. I did not need that anywhere. So, I can instead replace this rice and dal and so on by variables. I could have called them x_1, x_2, x_3 and y_1, y_2, y_3 and the whole thing would have worked out in the same way. So effectively, what we are saying by linearity is that if you have $c(\alpha(x_1, x_2, x_3) + (y_1, y_2, y_3)) = \alpha\, c(x_1, x_2, x_3) + c(y_1, y_2, y_3)$. That is what we mean by linearity.

This is a very important property, and this is very special to the kind of functions that we are dealing with, they are linear functions. So that is what this video was about. Remember, linear mapping. These are examples of linear mappings. Now, the claim is this can be, we could have done this in one shot by using the matrix. How is that? Because remember that, so let me just take here.

So, I can write this as the matrix times the vector x_1, x $\alpha(x_1, x_2, x_3) + (y_1, y_2, y_3)$. But we know that matrix multiplication commutes with the addition. So, I can instead write this as the matrix, the cost matrix that we had in the earlier slide, times x_1, x_2, x_3, and I can pull the alpha out because of constants remember come out plus the same matrix times y_1, y_2, y_3.

And then, of course, this I know is alpha times $c(x_1, x_2, x_3)$. I mean it is the same as the expression on top. It is equal to this expression. This this part is here and the second part is corresponding to $c(y_1, y_2, y_3)$. So, this is what we mean by a linearity.

(Refer Slide Time: 13:42)

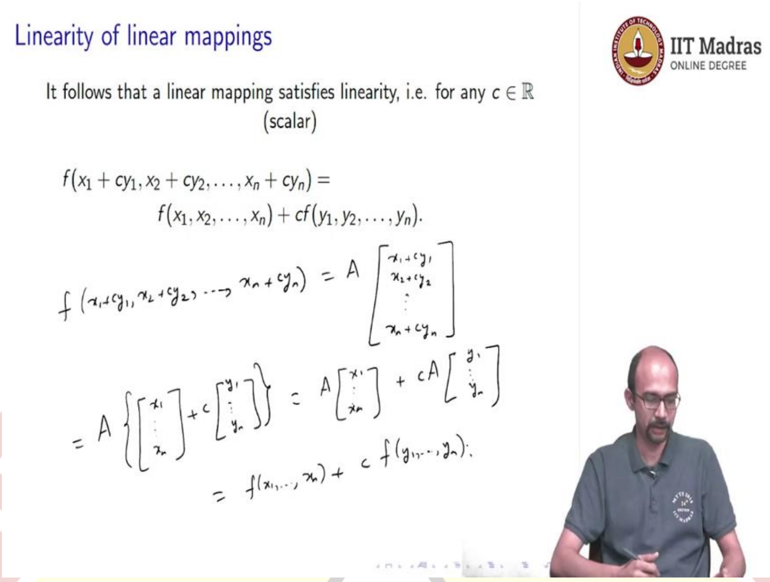


So now, we can talk in general, what is a linear mapping? So, a linear mapping f from $\mathbb{R}^n$ to $\mathbb{R}^m$. And now, we are not doing $\mathbb{R}^3$ or $\mathbb{R}$ we will just do it for any arbitrary n and m. mean something of this form where you have $f(x_1, x_2, \ldots, x_n) = \left(\sum_{j=1}^{n} a_{1j}x_j, \sum_{j=1}^{n} a_{2j}x_j, \ldots, \sum_{j=1}^{n} a_{mj}x_j\right)$, what are these a_{ij}'s, they are real numbers.

So, basically a linear mapping can be thought of as a combination or a, sorry, a collection of linear combinations. Except you arrange them in a vector, so it gives you a function. So linear mapping is a function, which has this form. So just to be clear, if you take something like $f(x) = x^2$, that is not a linear mapping or $f(x) = \log(x)$ that is not a linear mapping or $f(x) = e^x$ that is not a linear mapping. So, it should be linear, meaning it has this is very nice form.

So, we can of course write this as matrix multiplication, because notice that the right-hand side, if we think of it, instead of a row vector, if we think of it as a column vector then we can obtain it as multiplying $A * x$. What is A? A is exactly the matrix given by putting these coefficients into the appropriate places. So, the matrix the ijth entry of A is a_{ij} where a_{ij}'s is the corresponding coefficient above.

(Refer Slide Time: 15:19)



So finally, let us talk about linearity of linear mappings. This is exactly what we did for the cost vector. So, it follows that a linear mapping satisfies linearity, which means $f(x_1 + cy_1,\ x_2 + cy_2, \ldots, x_n + cy_n) = f(x_1, x_2, \ldots, x_n) + cf(y_1, y_2, \ldots, y_n)$.

Why is that? So, the reason is because I can write $f(x_1 + cy_1,\ x_2 + cy_2, \ldots, x_n + cy_n) = A\begin{bmatrix} x_1 + cy_1 \\ x_2 + cy_2 \\ \vdots \\ x_n + cy_n \end{bmatrix}$. So, here of course, I am, when I write it as A times just, I get a column letter, but you think of it as a row vector. So, when I say equal, I mean, you take the row vector corresponding to this column vector.

So, the second I mean the column vector over there $x_1 + cy_1$ I can break up into two separate vectors along with a constant. And then because matrix multiplication is commutes with or distributes over addition, we get this, which is exactly $A\left\{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + c\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}\right\} = A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + cA\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = f(x_1, x_2, \ldots, x_n) + cf(y_1, y_2, \ldots, y_n)$.. You could do this directly also, by the way, and I will encourage you to check this from the expression for the linear combinations.

So, just to recall what everything that we have done in this video. We saw the examples of grocery shops, first of a single grocery shop and the corresponding cost vector, and then of a bunch of grocery shops. And we saw that it was useful to arrange this as a function. We can treat the cost as a function, it is a linear combination, and that is the expression for the function, and it satisfies what is called linearity.

And then if we have several costs, which is what we had four different jobs, we arrange them into a cost function of which is a function of three variables. Taking values, depending on how many shops there are. So, it could be in, so we had three shops, so it took values in $\mathbb{R}^3$ so it is a function from $\mathbb{R}^3$ to $\mathbb{R}^3$, but, each coordinate of the expression that we had was a linear combination of the $x_1, x_2, \ldots, x_n$ of x_1, x_2, x_3. So, such a thing is exactly what is called a linear mapping. And then we observed that linear mappings are satisfied linearity. Thank you.