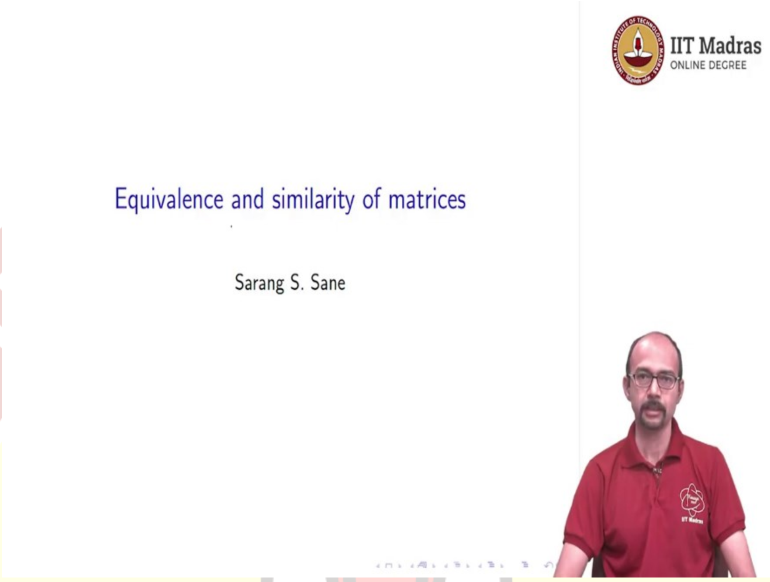
IIT Madras

ONLINE DEGREE

Mathematics for Data Science 2
Professor Sarang S. Sane
Department of Mathematics,
Indian Institute of Technology Madras
Equivalence and similarity of matrices

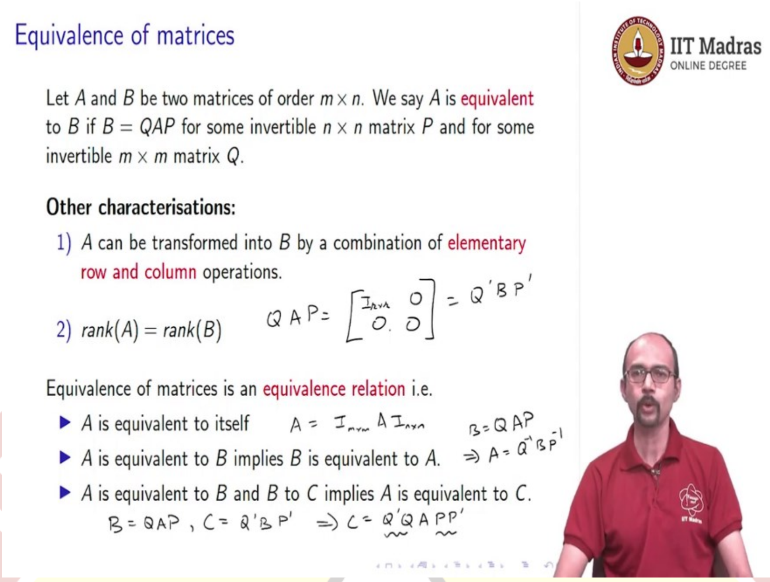
(Refer Slide Time: 00:14)



Hello, and welcome to the match 2 comp1nt of the online degree course on data science. In this video, we are going to talk about equivalence and similarity of matrices. So, we want to impose certain equivalence or what is called an equivalence relation based on two different relations, we are going to study this on the set of matrices.

So, the first 1 is called equivalence, this is the set of m by n matrices. And then the second 1 is for square matrices. So, the reason we want to do this is because there are certain what I will call invariants. This may not make sense right now, but that is fine called Eigen values, that you are going to see later on and these will come up not in this course, but in the data science part that you will see in the subsequent semesters where you will use these for certain calculations. So, without bothering about what everything that I said before, whether or not it made sense, let us get into the integrities of what is equivalence and similar similarity for matrices.

(Refer Slide Time: 01:28)



So, let us consider two, we will first talk about equivalence of matrices. Let us consider two matrices of order m by n, both have the same order, remember. So, we say A is equivalent to B, if B = Q × A × B, so we can find matrices, Q and P, which are invertible. So, P has to be because it is invertible, because I impose the condition of being invertible. First of all, they have to be square matrices. And since we are seeing QAP that means A is remember m by n, that means P has to be n by n, and Q has to be m by m, so they have to be invertible matrices, and B has to be equal to Q × A × P.

So, you have, if we can find Q and P or P and Q, such that B = Q × A × P meaning when you compute Q × A × P, you get B, then we say A and B are equivalent, or A is equivalent to B. So, the other way of viewing this is by doing our good old row and column operations.

So, if you can perform a series of row operations on A, and then a series of column operations on A, so that you can then obtain B, then we will say that A = B. So, you take all the column operations on the right, so you get P and you take all the row operations on the left, and that is you get Q by multiplying successively that is.

So, we have seen this idea before. And I am just trying to motivate how to get Q and P, if you want to actually get it. But we will see a different way of getting them. The other characterization is that the rank is the same. So, the reason this works is because remember that if you allow both row and column operations then you can take any matrix and reduce it to the form where you have block matrices, where you have the identity matrix, 0,0,0. I will say that carefully in a second.

So, equivalence of matrices is an equivalence relation. So, what is an equivalence relation? So, I am not defining it, but what it means is, A is equivalent to itself. Remember that we say A = B if there exists P and Q, so that B = P× Q × A × P. So, A is equivalent to itself is saying that there exists some Q and P, so that A = Q × A × P.

Now, I think it is clear what Q and P should be. We will write it down in a second. So, A is equivalent to B implies B is equivalent to A. Again, because of invertibility of P and Q, this is clear and A = B and B to C implies A is equivalent to C. So, let us quickly write down how to do this. A is equivalent to itself because you can write A as identity m cross m × A × identity n cross n, of course, identity matrices are invertible.

A is equivalent to B implies B is equivalent A. Well, we know that A is equivalent to B that means B = P × sorry, Q × A × P, and that implies by multiplying by the inverses that A = Q inverse × B × P inverse. So, A is equivalent to B, that is what we started with that means B is QAP, that will imply that B is equivalent to A.

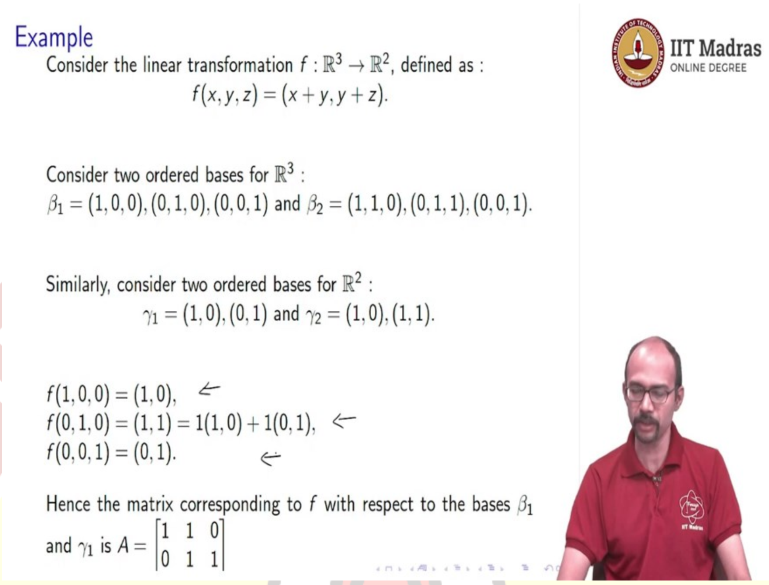
So, because we can write A as Q inverse × B × P inverse. So, that means you can multiply on the left by some invertible matrix and on the right by some invertible matrix to get A from B, so B is equivalent to A. And then A = B and B to C implies A is equivalent to C. This is because B = QAP, C = let us say, Q prime, B, P prime, then you can replace B by the first part. So, C = Q prime × Q × A × P × P prime.

And now Q and Q prime are invertible, that means this product is invertible. Similarly, P and P prime are invertible, so this product is invertible. So, I have an invertible matrix on the left and on the right, which I can multiply to A, in order to get C so that is why A is equivalent to C. And now we can just say A and B are equivalent because this is an equivalence relation.

So, once you know that A is equivalent to B, B is also equivalent to A, so we can just say A and B are equivalent. So, from now on, I may some× say A and B are equivalent matrices, that means A is equivalent to B or B = A both are the same. Now, let me come back to this rank of A is rank of B.

What I meant here was that you can get Q and P so that QAP is of this form. And you can also get Q prime and P prime so that Q prime B, P prime is of this form. And then that is why A is equivalent to this matrix here if we call it C, and B = this matrix C, but this is an equivalence relation, that means A is equivalent to B. So, that is 1 way of checking that if rank of A is rank of B, then indeed, they are equivalent. Conversely, if they are equivalent then I will leave it to you to check that indeed, the ranks are the same that is a little bit of manipulation of column and row spaces. So, we know what it means to say that two matrices are equivalent.

(Refer Slide Time: 07:48)

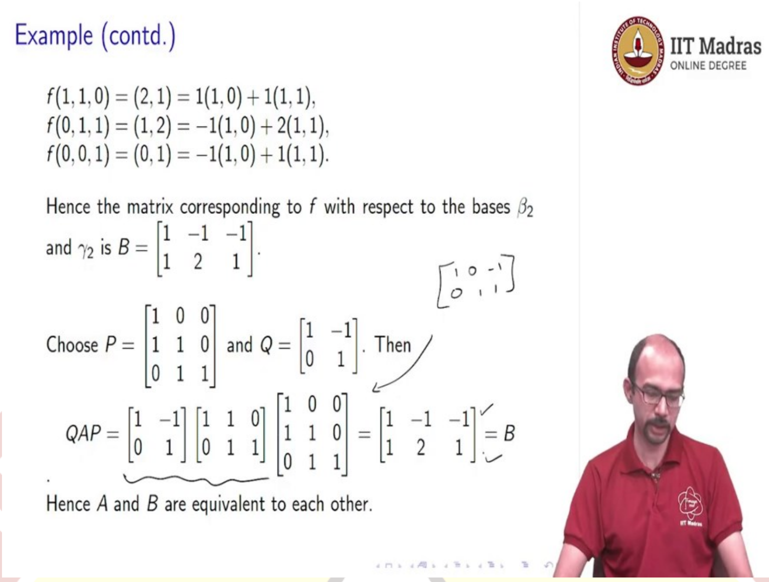


So, let us look at the linear transformation f from $\mathbb{R}^3$ to $\mathbb{R}^2$, given by $f(x, y, z)$ is $x + y$, and $y + z$. So, let us consider two ordered bases for $\mathbb{R}^3$, the standard ordered basis, and another ordered bases which is 1, 1, 0, 0, 1, 1 and 0, 0, 1. And let us consider two ordered bases for $\mathbb{R}^2$, which is a standard ordered bases and the bases, 1, 0 and 1,1.

So, let us express the standard ordered bases for R3 in terms of the standard ordered basis for R2 after applying f. So, if you apply f to 1, 0, 0 you get 1 + 0, , 0 + 0, so that gives you 1, 0. If you apply f to 0, 1, 0 you get 0 + 1 and 1 + 0, which is 1, 1 and then you write it in terms of the standard bases, which is 1 × 1, 0 + 1 × 0, 1. And then if you apply f to 0, 0, 1 you get 0 + 0 and 0 + 1, which is 0, 1 and that is just 1 × the second vector in the ordered standard bases.

So, if you want to write down the corresponding matrix for this linear transformation, that is corresponding to the ordered bases, standard ordered bases for $\mathbb{R}^3$ and standard, ordered basis for $\mathbb{R}^2$ you get 1,0, 1,1 and 0,1. So where is this 1,0 coming from it is coming from here. Where is this 1,1 the second column coming from, it is coming from these coefficients. And the third 1 is coming from these coefficients. This is something we did in our previous videos.

(Refer Slide Time: 09:30)



Let us do the same thing for the other two bases. So, the other standard, the other ordered bases for $\mathbb{R}^3$ and the other ordered bases for $\mathbb{R}^2$. So, if you look at $f(1,1,0)$ that is 1 + 1 , 1 + 0, that is 2,1, which you can express as 1 × 1,0 + 1 × 1,1. In these the ordered bases the second ordered bases γ 2.

Similarly, if you take $f(0,1,1)$ this is 0 + 1 , 1 + 1, which is 1 , 2. And for this, if you order, if you write it in terms of the ordered basis it gives you - 1 × 1 , 0 + two × 1 , 1. And for the last 1, you have f of 0,0,1, which is 0 + 0 , 0 + 1, which is 0,1, which is - 1 × 1,0, + 1 × 1,1.

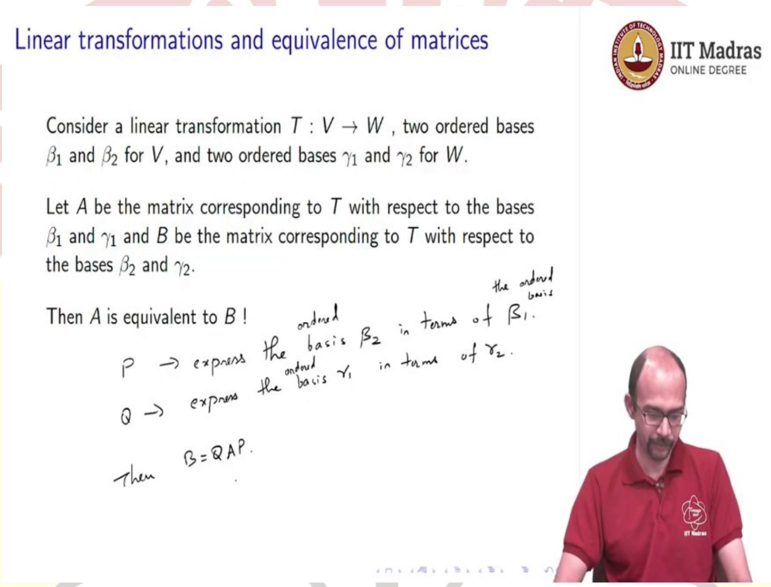
So, if you look at the corresponding matrix, first column is going to come from the first equation coefficients, the second column is going to come from the coefficients of the second equation and the third column is going to come from the coefficients of the third equation. So, it will look like 1,1 - 1,2 and - 1,1.

So, what have we d1 so far, we have taken a linear transformation and we have written down two matrices corresponding to two different ordered sets of ordered bases. So, choose P to be this matrix 1, 0, 0 and 1, 1, 0, 0 ,1, 1 and Q to be 1 - 1,0,1. So, I, this is something I produced out of the air and here is what happens. So then, Q × A × P, this is something you have to compute, I have d1 the computation here.

This is 1 - 1, 0, 1 × 1, 1, 0, 0, 1, 1. So, this was A × 1, 0, 0, 1, 1, 0, 0, 1, 1. So, if you multiply this out, let us see what you get after doing the first multiplication, you get the first row in out of out of these two is, so 1 - 1 × so this is 1, then this is 0, then this is - 1, and then you have 0,1,1.

And you want to multiply this to this matrix. So, if you do that, let us see what we get. So, if you do 1, 0 - 1, indeed, you get 1, then 1, 0 - 1 × 0, 1, 1 that is - 1, and similarly a - 1. So, this first row is indeed achieved and then for the second row, you take 0, 1, 1 and multiply it over there so that gives you 1, 2 and 1, so indeed the second row is achieved. So, if you do this multiplication yourself, you will be able to see that Q × A × P is indeed equal to B. So, these matrices are equivalent to each other. So, this is not some a magic.

(Refer Slide Time: 12:49)

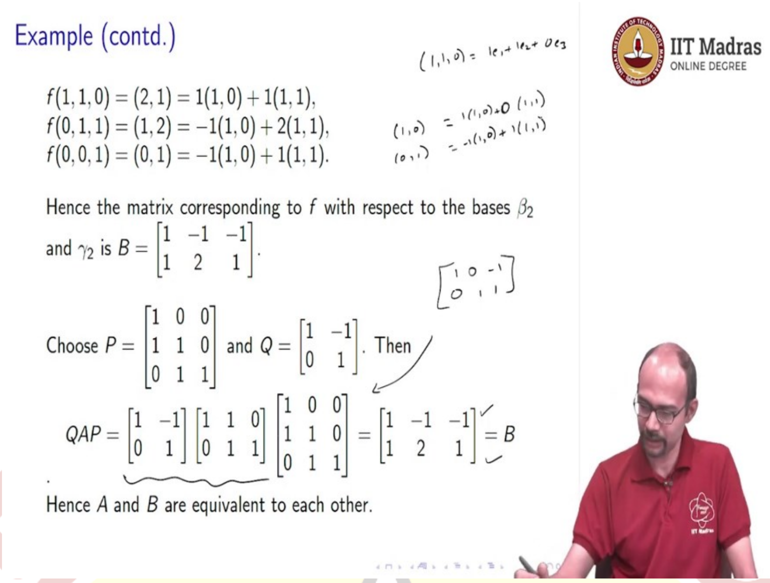


So, consider a linear transformation T, V to W, and two ordered bases β 1 and β 2 for V, and two ordered bases γ 1 and γ 2 for W. So, like in the previous example, where we had V was R3, W was R2 and β 1 and γ 1 with a corresponding standard ordered bases β 2 and γ 2 are some other fixed ordered basis.

So, we took A to be the matrix corresponding to β_1 and γ_1. So, similarly, here, you take A to be the matrix corresponding to T with respect to the basis β_1 γ_1 and you take B to be the matrix corresponding to T with respect to the basis β_2 and γ_2, this is exactly what we did in the previous example, then A = B.

So, the previous example is an example of this phenomenon that is that we have menti1d here. So, the question is, how do I get Q and how do I get P? So, for P, express the bases β_2 in terms of β_1, and for Q express the bases, I should say ordered bases, everywhere we have ordered bases, the ordered bases γ_1 in terms of γ_2. So, this is how I got those two matrices. Let us quickly look at that example again.

(Refer Slide Time: 14:46)

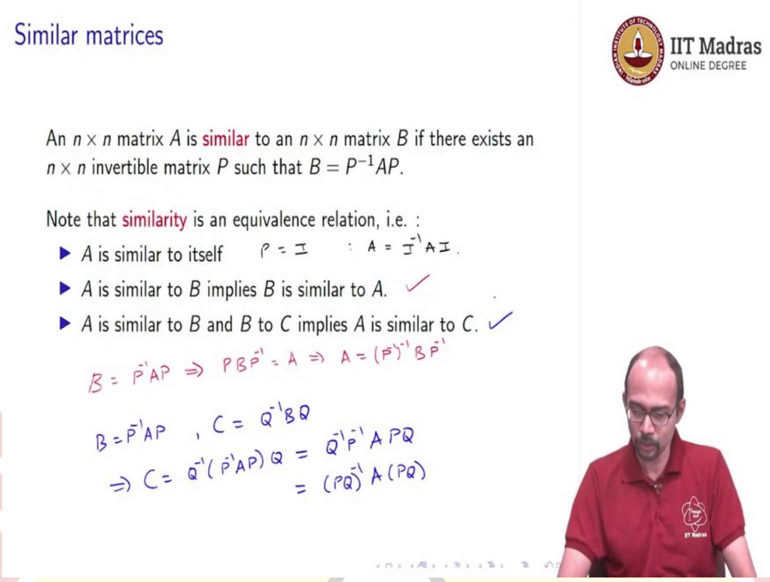


So, right. So, let us look at this example again. So, if you look at the matrix P, the matrix P is exactly expressing this basis here 1, 1, 0 so this is 1 × E_1 + E_2 + 0 × E_3. So, 1, 1, 0 goes into your first column, and similarly each of these, so you have 0, 1, 1 which goes into your second column and 0, 0, 1 which goes into your third column, and so that is how I get P.

And for Q, what I did was, I wrote 1, 0 and 0, 1 in terms of 1, 1 and sorry, in terms of 1, 0 and 1, 1 so this is 1 × 1, 0 + 0 × 1, 1. That is why I get 1, 0 in my first column and 0, 1 is - 1 × 1, 0 + 1 × 1, 1 that is how I get - 1,1.

So, this is a this is a strategy for getting P and Q and this works for any general linear transformation and these fixed ordered basis β 1 β 2 and γ 1, γ 2. So, yeah, so to complete this express this, get P and Q from there and then you will find that B = QAP. So, this is something you have to check. I will leave it at that since this is not really a proof-based course.

(Refer Slide Time: 16:32)



So, we have exhausted equivalence, let us look at similarity of matrices. So now, we are going to restrict ourselves to square matrices. So, we will say that two square matrices of size n by n are similar or rather we will say that A is similar to B. So, we have two n by n matrices A and B and we will say A is similar to B, if there is an invertible matrix P, it has to be of size n by n such that B = P inverse AP. So, now the role of Q is being played by P inverse. So, this is a very big change. So, if you can express B as P inverse AP, then we will say that A is similar to B.

So, again similarities and equivalence relation by which we mean A is similar to itself. A similar to B implies be similar to A and A similar to B and B to C implies A similar to C. Let us quickly see why this is the case. So, for the first 1, again, you can take P to be identity, then A = identity inverse A identity. Well, identity inversus identity, so this is just identity × A × identity, which is certainly A. For the second 1. Let us see how to do the second 1, these are on the same lines that we have d1 the previous part.

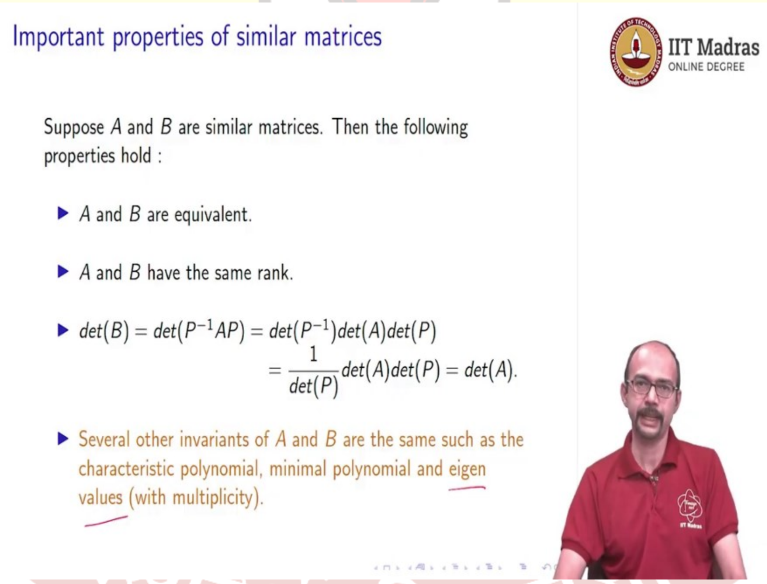
So, A similar to B means B = P inverse AP, but that means, because P is invertible, I can multiply by P inverse on the right and P on the left, so I get PBP inverse = A. So, just rewriting that P = P ×, so I want to write this as P inverse, inverse. So, why did I write it like that, because I want to express this in terms of the inverse of whatever is occurring on the right term, P inverse × B × P inverse.

Now, you take your P inverse, call it something else call it P prime, if you want, then we have an invertible matrix P prime such that A = P prime inverse × B × P prime, which is exactly saying that B is similar to A. So, the second 1 is satisfied. And the third 1 again on the same lines as we did our previous part. So, let us see.

So, we have these B is P inverse AP and C is let us say Q inverse AQ, these are both n by n invertible matrices P and Q. So, then C = my bad, this is not B, but this is not A but this is B. So, C = Q inverse and now you put in the values put in the first equation, which is B is B inverse AP into the second, then you get Q inverse P inverse × A × PQ, but remember Q inverse P inverse is exactly PQ inverse. And now if I call P × Q by something else, let us say P prime, then this is saying that C = P prime inverse × A × P prime, which says that A is similar to C.

So, this proves the three things. So, this says that similarity is what is called an equivalence relation. So, now I can say, especially because of the second 1, instead of saying that A is similar to B I can just say A and B are similar because it does not matter which is similar to which the other 1 always follows. So, A is similar to B, can be replaced by saying A and B are similar matrices.

(Refer Slide Time: 20:14)



So, why do we really care about similar matrices? So, the reason we care about similar matrices is because of some of these properties. The first is that A and B are equivalent, this is clear, because equivalent means that there is P and Q so that B is Q × A × P, where P and Q are invertible, but if they are similar then you can take Q to be P inverse, and P to be P. So, this is clear. A and B have the same rank as a result, the determinant is the same. Now, this is a big difference.

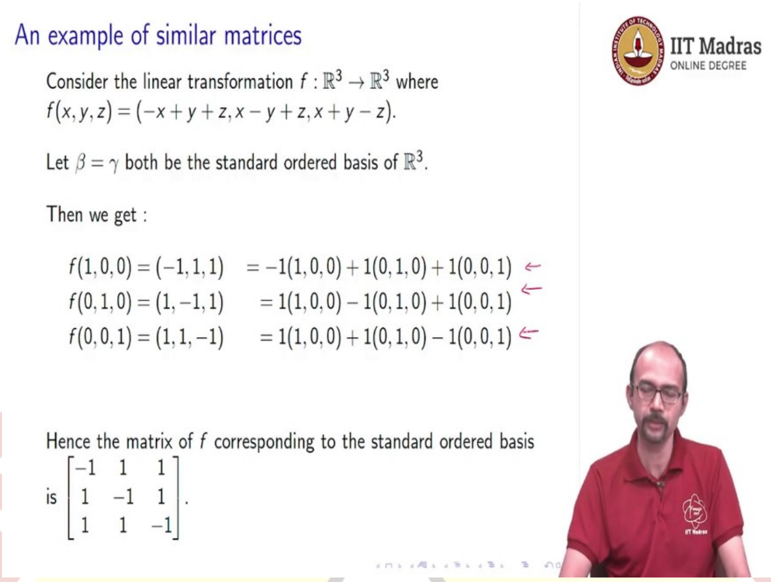
Remember, now we are in square matrices, we could not have made any such kind of statement before. The determinant of B is determinant of P inverse × AP, but determinants multiply, if you remember, then that means this is determinant of P inverse × determinant of A × determinant of P. But now, although, matrices do not commute real numbers do, and determinants give you real numbers.

So, I can write this as 1 by determinant of P, because determinant of P inverse is 1 by determinant of P × determinant of A × determinant of P, and then because these are real numbers, I can shift 1 by determinant of P to the right, and then determinant of P divided by determinant of P gives me 1, so I get just determinant of A. So, they have the same determinants. This is a very useful property and this is 1 of the properties that you will be using later on. Not in this course, but in a different course for certain reasons.

So, the reason, in fact, is that there are other invariants, so 1 such as called the Eigen values which you are going to study later. So, I have written some names here. And this, you are not expected to make sense of this, but it is good, it will be good if you remember these, this just as a statement without understanding when you do your next course.

So, several other invariants of A and B are the same such as the characteristic polynomial, the minimal polynomial and the Eigen values with multiplicity. We are not going to define this right now, but this is just as a feeder into the next course, where you will be using some of these things, especially, this thing called the Eigen values. So, just remember this name. Does not matter if you do not understand what it is now. So, this is why we are studying this idea of similarity.

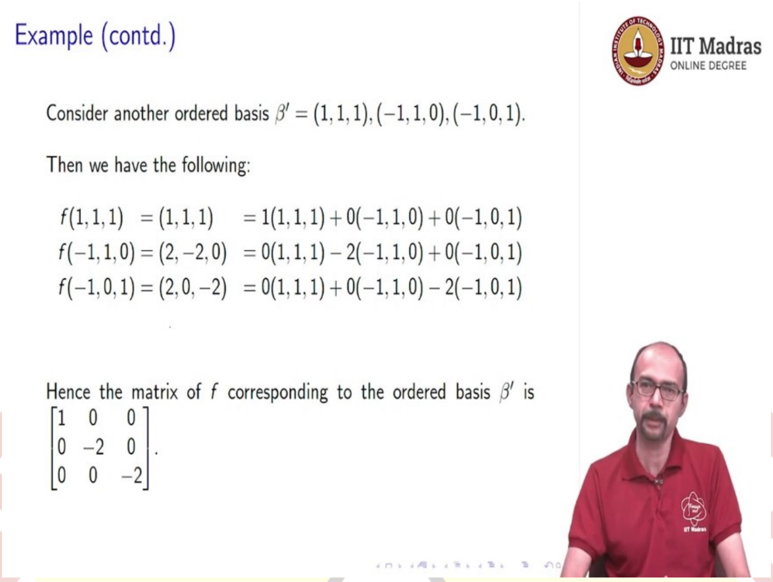
(Refer Slide Time: 22:47)



Let us do some examples or maybe 1 example. So, suppose you have a linear transformation f of x, y, z is $-x + y + z, x - y + z$ and $x + y - z$. Now, we are going to use the same standard the same bases. Earlier, we had two different bases. So, we had β_1 and γ_1, and then we had β_2 and γ_2. But here, remember that this linear transformation is from $\mathbb{R}^3$ to $\mathbb{R}^3$ itself. So, we are going to use the same ordered bases. So, let us start with β being γ being the standard ordered basis.

So, let us do what we did before. Let us express f of 1,0,0, f of 0,1,0, and f of ,0,0,1 in terms of the standard basis. If you do that, you get f(1,0,0) is - 1,1,1, f of 0,1,0 is 1 - 1,1 and f of 0,0,1 is 1,1 - 1. So that means these things exactly end up being the coefficients, when you write it in terms of the standard ordered bases. And as a result, the matrix of f corresponding to these, the standard ordered bases is coming from these give you your first column these coefficients give you're your second column and these coefficients give you your third column.

(Refer Slide Time: 24:13)



Let us choose some other bases, β prime and γ prime is the same. So, this is the same, so this is β prime, and it is also γ prime. So, let us express this f of each of these vectors in terms of as linear combinations of these vectors themselves. So, if I do f of 1,1,1, I get back 1,1,1. Let us remember what the linear transformation was it was $-x + y + z, x - y + z$ and $x + y - z$.

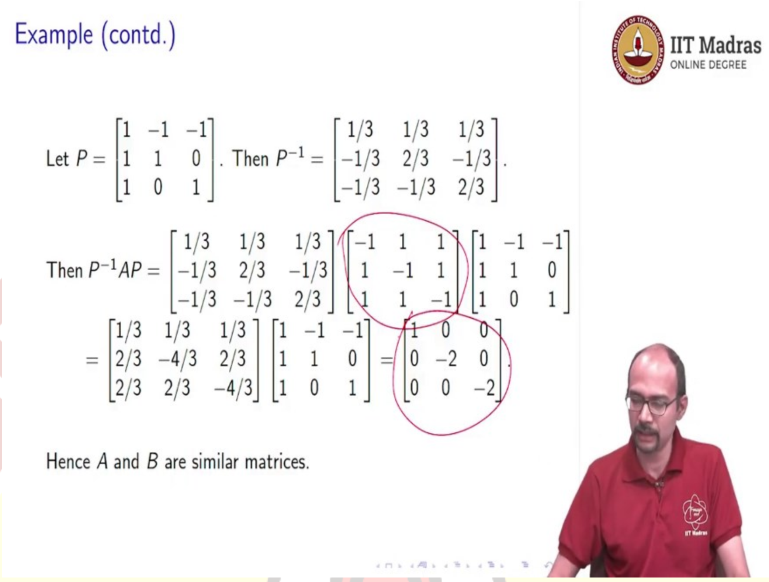
So, in that case we get f of 1,1,1 is - 1 + 1 + 1, 1 - 1 + 1, 1 + 1 - 1 which is 1,1,1, so which is the first factor itself. So, this is 1 × 1,1,1 + 0 for the other 0 is in the coefficients for the other vectors. f of - 1,1,0 is - of - 1, which is 1, + 1, + 0, which is 2, then - 1 - 1 + 0, which is - 2, and then 1 - 1, rather - 1 + 1 + 0, which is 0, so you get 2 - 2, 0 and f of - 1,0 , 1 is, again, you will, if you do it you will get 2,0 - 2.

And now if you write these in terms of the vectors that you have, the second 2 - 2,0 is just two × the second vector or other - two × the second vector. And the f of the third vector is - two × the third vector. So, what we have got is, there is a basis here, which is β prime, so that when I apply f on that basis for these particular vectors in the bases, they just scale.

So, f of 1,1,1 is 1,1,1, f of - 1,1,0 is - 2 × - 1,1,0 and 4 - 1,0,1 is - 2× - 1,0,1. So, for these particular bases, the linear transformation behaves very nicely, it just scales. 1 of the vectors remains the same so it scales by 1, the second and third vector get scaled by - 2.

So, the corresponding matrix is diagonal matrix 1, 0, 0, 0 - 2, 0 and 0, .0 - 2. And now, we know that diagonal matrices are particularly easy to understand. So, what we see here is that, we have the same linear transformation in terms of 1 basis the standard ordered bases, we got a matrix which was not hard, but not maybe not so easy. But in terms of this bases, this new basis β prime ordered bases, we saw that it is a diagonal matrix.

(Refer Slide Time: 27:17)

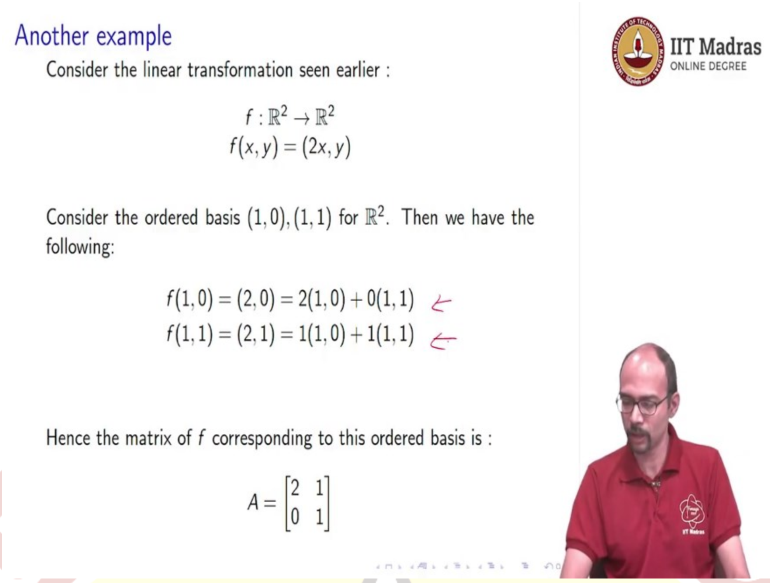


Now, let P be 1 - 1 - 1,1,1,0,1,0,1 compute P inverse, this is something you can do, I have d1 it here. It is 1 third, 1 third, 1 third, - 1 third, two third - 1 third, - 1 third, - 1 third, two third. And now, if you multiply, so, you look at P inverse AP, then you get. If you multiply this carefully, I have d1 it here. So, I will suggest you check that this calculation is correct, you get the matrix B.

So, this last matrix is exactly your diagonal matrix B. So, you get that A and B are similar matrices. So, what have we seen, we have seen exactly what happened in equivalence, the same kind of thing is happening for similarity, but this time, instead of having different bases, your β and γ are the same ordered bases because your vector space V and is the same vector space, it is V and V, so you have a linear transformations from V to V.

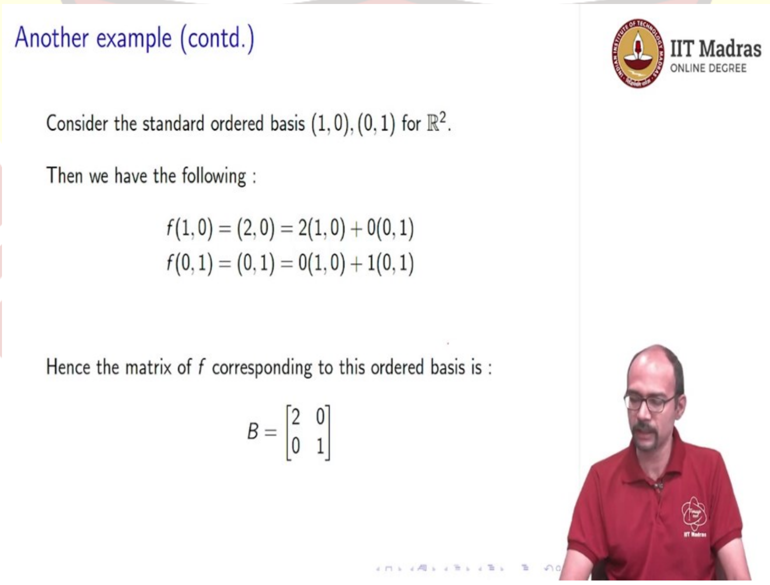
So, here we have $\mathbb{R}^3$ to $\mathbb{R}^3$. So, you take the bases, β and β. In our case, we do identity, sorry, the standard ordered bases and you got the matrix A, which is this matrix here. And in terms of a different ordered basis, which we call β prime, where implicitly γ prime was equal to β prime, we got this matrix. So, remember that now the basis that you are choosing on both sides are the same. So, the β and γ have to match. And then if you write down these matrices, they end up being similar matrices.

(Refer Slide Time: 28:58)



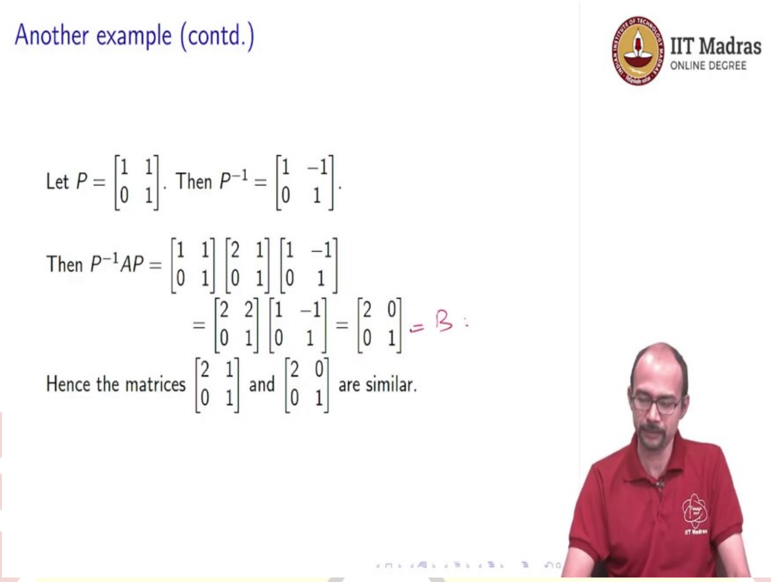
So, this is a general phenomenon. Let us look at another quick example. So, f of x, y is $2x\,, y$. So, if you look at the ordered bases 1,0 and 1,1 you can look at. So, f of 1,0 is just 2,0, f of 1,1 is 2,1 and you can write this in terms of our bases that you have 1,0 and 1,1. So, these bases is if you want β and γ both. And as a result, the corresponding matrix is the coefficients coming from here, coefficients which go into the first column, coefficients from there, which go into the first, second column, so it is $\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$.

(Refer Slide Time: 29:37)



Instead, you can look at the standard ordered basis. So, this is this time your β prime and γ prime, and it is clear it is $2x\,, y$. So, f of 1,0 is 2,0 f of 0,0,1 is 0,1, so this is scaling your first vector by 2 and it is returning your second vector as it is. So, this gives us a diagonal matrix 2,0,0,1.

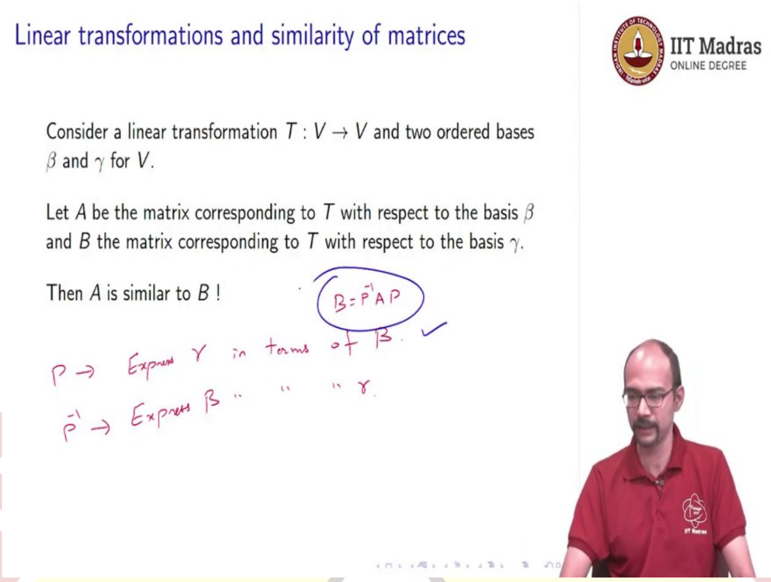
(Refer Slide Time: 30:01)



So let P be 1,1,0,1 P inverse is 1 - 1,0 you can compute this. And now if you compute P inverse AP, this is a very simple calculation, which I have d1 here, you get your diagonal matrix, which is exactly the matrix B. So, this is exactly B, so the matrices 2,1,0,1 and 2,0,0,1 are similar.

So, we have seen two examples where we took linear transformations from two vector spaces to themselves expressed each time these linear transformations as basis with the same basis on both sides, and then we change the basis. Again, with the new basis on both sides, we computed the matrix and we saw that these matrices are similar.

(Refer Slide Time: 30:51)

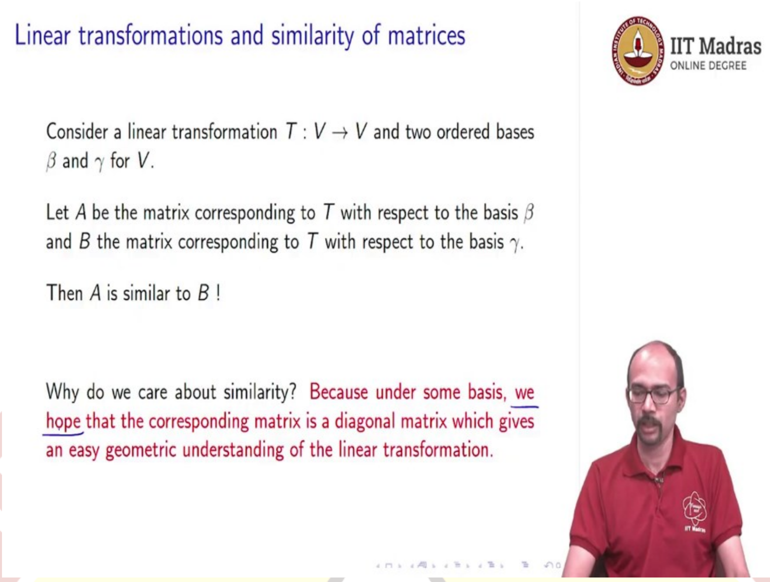


So, this is a general phenomenon, and the proof for this is exactly the same. So, consider a linear transformation T, V to V and two ordered bases β and γ for V. Let A be the corresponding matrix with respect to β and B be the corresponding matrix with respect to γ, then A is similar to B. So, that means I have to produce P. How do I produce P?

To produce P you write, so express, so we want that B is P inverse AP, so express, so for P you express the, you express γ in terms of β, and P inverse is well you can get it from P by inverting or you can get it from B by expressing β in terms of γ. What does this mean? This means you take each in an ordered way, you take. So, this first 1 means you take the first vector of γ, write it as a linear combination of elements of the vectors from β, take the second vector of γ write it as a linear combination of β and so on, so you get that.

So, if you have n many, you write all those n many things, and then you get a matrix. For the first equation that you have put the coefficients into the first column, for the second equation, put them into the second column and so on for the nth equation, put them into the nth column, that is your matrix P. Same thing, but the opposite way, that is your matrix P inverse or you can get it just by inverting P. And you will find that you end up with B is P inverse AP. So, again, it is not a proof-based course, I would not prove this fact, but this is how you get P fine.

(Refer Slide Time: 32:49)



Why do we care about similarity? I mean, this is maybe the most important question. So, the reason is that, as we saw in the previous two examples, sometimes. So, we hope, so, the underlined part is we hope, because under some basis, we hope that the corresponding matrix is a diagonal matrix, which gives an easy geometric understanding of the linear transformation.

So, if you have a diagonal matrix, as I said it saying that some particular vector scales, some other particular vector scales and some third vector letting our three scales and give me how much it scales by and these three form a basis. So, that is the intuitive picture you have in mind.

So, let us quickly recall what we have d1 this video. We have seen the notion of similarity and equivalence of matrices. Equivalence was, both of these are equivalence relations. Equivalence is for rectangular matrices, I mean, which means matrices which need not be square. So, we say they are equivalent to B is QAP, then A and B are equivalent.

For similarity, we say that B is P inverse AP, I mean, if these B is P inverse AP then we say A and B are similar. And these notions are you are useful because they sort of help you to understand how to go between different basis on the vector space if you want if you have a particular linear transformation and you choose different basis, how do you go between the corresponding matrices that is really what equivalence and similarities telling you.

And similarity is particularly important because, it may so, happen that your linear transformation may not, the algebra that when you write it down, it may look like a very difficult entity, but under some particular basis, it may suddenly start looking very nice. This is really the main point behind why we study similar. So, thank you for your attention.