

MITOCW | 2. Multiplying and Factoring Matrices

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation, or to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

PROFESSOR: So, are we ready to go? Any questions on 18065? So it will be, as I said before, a mixture of linear algebra and math questions, along with online using the material. OK. So I'm, in this first week or two, reviewing the highlights of linear algebra. And I've reached this point, to remember-- well, so we-- I just said two words about multiplying matrices by using column times row as a way to do it.

And now, I want to illustrate that by the five key factorizations of matrices. OK. So what are they? And do you recognize them? Everybody uses those letters. In fact, some of those letters, like LU or QR, would be the most used MATLAB commands in linear algebra. So a A equal LU of, maybe-- say something I'll develop today-- but it's about elimination. Solving linear systems.

So that's always the start of a linear algebra course. But it will go fast here. I just want to show you a different way they get to L times U-- lower triangular times upper triangular. Probably you've seen that-- those triangular matrices. So do you know what QR is? What's QR about?

AUDIENCE: Least squares?

PROFESSOR: Least squares is the big application, the factorization. So what kind of a matrix gets that letter Q?

AUDIENCE: Orthogonal?

PROFESSOR: Orthogonal. The columns are orthogonal. Often orthonormal. So orthogonal means they're perpendicular to each other. And orthonormal means they're unit vectors. So that is-- so Q often represents a matrix with orthonormal columns. So you-- we could say Gram-Schmidt, if you want to remember a couple of old timers whose algorithm produces Q and R.

How about this one? This is really a central one in math-- pure math, applied math, everywhere-- applications. So S stands for symmetric. So this is a special factorization for symmetric matrices. And you can see that it's symmetric. This lambda is the diagonal

eigenvalue matrix-- always λ for eigenvalues. Q is like that Q -- different Q , of course. That Q you can find just straight forward from Gram-Schmidt.

This Q has the eigenvectors. So you don't find eigenvectors without some extra work. OK. So that's eigenvectors. Yeah. So that would be worth expanding. So here are the eigenvectors of the matrix S -- n of them, normalized. Here are the eigenvalues λ_1 to λ_n . And here are the eigenvectors now transposed.

So remind me of the great fact about-- two facts, I guess-- one fact about the eigenvalues and one fact about eigenvectors. This is an important fact statement in linear algebra. What do we know about the eigenvectors? Oh well, I guess I've given it away. The eigenvectors are orthogonal. That's very important. Makes some matrices-- well, they're beautiful matrices. They're the kings of linear algebra. Q s are the queens, in my opinion.

Orthogonal matrices are the queens, and symmetric matrices are the kings. So these are orthonormal eigenvectors. And the key point-- an important point that's implicit here-- is, there are n of them. There is a complete set. The matrix can be diagonalized. And those-- well, what's special about the eigenvalues?

Other matrices could be $Q \lambda Q^T$. But symmetric matrices are something additional about λ .

AUDIENCE: They're all real?

PROFESSOR: They're all real. So eigenvalues are real. And eigenvectors are orthonormal-- can be chosen orthonormal-- can be chosen, I guess I have to say. OK. Good. Good. Good. Oh, now maybe I'll use that as an example of matrix multiplication. So let me just do that here. Simple matrix multiplication, but it makes the point. So $Q \lambda Q^T$. OK. Well, what was my point about matrix multiplication? Let me-- it really involved two matrices. Here, I unfortunately have three.

So I'm going to have to squeeze λ in with one of the Q s, to see it nicely as two matrices. Shall I just do that? Yeah. Now I've made it two matrices. That was easy. OK. Now what's the rule? In the first notes, this was A and this was B . And when you multiply two matrices, the rule is, this is columns of $Q \lambda$ times rows of Q^T . I'm multiplying columns by rows.

And so it's a column vector times a row vector, and that gives us a matrix. So each-- and it's a

special matrix. So this is our column. This is a row. And when I multiply n by 1 times 1 by n , I get an n by n matrix. And it's pretty special. And what is the special fact about I'm sort of recalling from last time. What's special about a column times a row? It's rank is special. It's rank is 1 .

It's column space-- well, the only column around is this one. So all columns are multiples of this guy. All rows are multiples of this guy, as we could see from an example. Shall I just do an example? $1, 2$ times $3, 4$, to take a random example. So that would give us $3, 4, 6, 8$. And sure enough, the columns are multiples of $1, 2$. The rows are multiples of $3, 4$. And the rank is 1 .

OK. So those are the building blocks. Now I want to build something. So here we go. So this is a sum of rank 1 . Sum of rank 1 . Sum of column times row. So I take column 1 times row 1 . That's my first thing in the sum. So column 1 of that. So you see I had to sneak the λ to have just two factors. So what's column 1 of $Q \lambda$? That's a good question.

So column 1 of Q is q_1 -- the first eigenvector. But now, I am multiplying by this diagonal matrix. Do you see in your mind what's the first column of $Q \lambda$? Just think about that a second. So here's-- we can steal any little corner for a matrix. So here's q_1 , and the rest of the columns. And then here is λ_1 . And that's $Q \lambda$. So I'm putting those together. And I'm asking, what's the first column of the answer? Can you see how that works?

AUDIENCE: $q_1 \lambda_1$?

PROFESSOR: It's... sorry?.

AUDIENCE: $q_1 \lambda_1$?

PROFESSOR: $q_1 \lambda_1$. Exactly. That λ_1 will multiply q_1 . These other λ s multiplied later columns. So the first column is λ_1 times q_1 . Right. So it's the first guy, λ_1, q_1 . And then the first row of this will be q_1 transpose. That's the first guy in our sum of n things. And let me put the next one and the last one. λ_2, q_2, q_2 transpose. And λ_n, q_n, q_n transpose. That's really a nice right way to write to breakup the product, $q \lambda q$ transpose.

This is called the spectral theorem. So that's the symmetric-- that's S . That's S there, is that. So that's-- we've broken up S into rank 1 pieces. That's like a constant theme. And these rank 1 pieces are quite special because they're symmetric. $Q_1 q_1$ transpose will be symmetric. And oh, so can we-- so let's just-- I follow the rule for multiplying matrices. But maybe I could just

check that it's the right thing-- that it came out right.

So what do I mean by checking? I guess I'll just check about S times q_1 . So look at S -- this thing-- times the first eigenvector, and what do I get? OK. So you'll like this. I've split up S into a sum of rank 1 pieces. And that splitting is-- you see it all over. It's really showing you what the pieces of the symmetric matrix are. And now I'm just going to check that that's a correct formula for S , so I'll multiply it by q_1 . And I'm hoping to get the right thing.

And what do I actually get? If I multiply this whole business times q_1 , I get λ_1, q_1, q_1 transpose-- that's the first guy times my q_1 -- plus-- right? I'm multiplying S by q_1 , and this first term gave me that. And what does the next term give me? Put me out of my misery here. I'm looking for this thing to simplify like mad. OK. So what's the second term?

When I multiply this guy by q_1 , what do I get?

AUDIENCE: Zero.

PROFESSOR: Zero. That's right. That's what we want. And when I multiply the last guy by q_1 , I get zero, because the q s are orthogonal. So this is all I get. And then-- so, I don't need this plus anymore. That's it. And then what can I do to improve that little somewhat repetitive formula for the answer? What do I want to do finally? I want to remember that the q s are normalized. They're unit vectors. So what does that tell me here?

AUDIENCE: Q_1 transpose.

PROFESSOR: Q_1 transpose times q_1 . This is just 1. It's-- that's what normalized means-- that the length squared is the length-- the length of the vector squared. And it's 1. So I can cancel that term. And I'm getting the right answer. That's all-- that's all this was about. I was just checking and wanted to see how it would fall out. And it falls right out that this formula is the correct matrix S , because it's got the right eigenvectors, q s.

And it's got the right eigenvalues λ s. So it's gotta be the right matrix S . Is that OK? That's like a first example to see how this-- splitting into rank 1s-- gives you back what you expect easily enough. It gives you the information you expect. OK. So that's the symmetric eigenvalue picture for symmetric matrices. And we'll see it again.

It's-- well, all five of these are big, are important. I don't know if you know this one, but it's going to be a foundational factorization for this course and for all of data science. Do you know

its name? So, what does it mean, first of all? Just a comment on this, and then we'll save it for a couple of weeks.

So this view is actually an orthogonal matrix. And so is V . So it has two orthogonal matrices. So that's why people call them U and V rather than Q_1 and Q_2 , which was too much to get subscript. So orthogonal times diagonal times orthogonal. And we'd say, orthogonal, diagonal, orthogonal. And 18.06 would, as of now, reach this topic because it's jumped up in importance. And it's called?

AUDIENCE: Singular value decomposition.

PROFESSOR: Singular value decomposition. Well, those are long words, so everybody calls it the SVD-- the singular value decomposition. The point is it works for every matrix-- rectangular matrices. There's no issue of, does it have enough eigenvectors or not? That's an issue here. Well, it's an issue here. Not every matrix has got enough eigenvectors to make that work.

Every matrix that one works because instead of one set of vectors it's got two matrices-- two different sets of singular vectors. Oh, we'll see that. That's important. OK. So that's really a quick overview of fundamental factorizations. And I'd like to say just another word about elimination, A equal LU , and then we'll leave it alone. So elimination. Yeah. Do you remember that first beginning of linear algebra, when you're solving ax equal b ? You do these row operations.

Can I just-- what I want to say is, all those row operations that you do are perfectly expressed by L times U . And so that's a key point in 18.06, but I have a different way to look at it. So that's what I wanted to show you. I have a-- I want to show you a sum of rank 1's, or row times column. It fits in today. So I'd just like to see, why does a matrix invertible-- this is a square matrix, now invertible. And it factors-- if all goes well with elimination and the pivots are non-zero-- it factors into lower triangular times upper triangular.

So that's a key step that is-- that MATLAB would do with a lu of A -- would produce those two factors. Now I want to do them in a column times row way, which I just realized late, was a neat way to do it. So can I take a matrix and do elimination? How big a matrix shall I take? 2 by 2 [INAUDIBLE]. 3 by 3 Somebody's not convinced totally by 2 by 2

Let me do it 2 by 2 and then if you want-- if you really want a 3 by 3 do it. OK. Here's 2 by 2, 4, 3, 7. How's that? OK. So what is-- yeah. So let's remember what elimination does. It

subtracts a multiple of that row from that one, to get to 2, 3. And the multiple is 2. So it knocks out the 4. Two 3's are 6. So it leaves a 1. And we're-- oh, yeah. Thanks for allowing me to do 2 by 2. I've already done it now. I've reached U. So here's A. And here's U-- the upper triangular guy with the pivots on the diagonal.

And then the question is, express that step in matrix language. And the right answer is I times U . So the right answer is that this A is-- so I'll just erase that letter U -- is I times U . So what is L ? L is the lower triangular guy. And it has there the number that you used here. And what was that? I subtracted 2 of that row from this. So I want a 2 there.

So that would be-- I would call that a multiplier. I multiplied row 1 by 2. Subtracted to get that 0. And for a 2 by 2 example, I was finished. OK. Now I want to see this-- so there is I times U . It happened. Right. I would like to see how I times U comes out of this row-- column times row. So let me start-- let me think again. So really the point of elimination-- why did we do this in the first place?

Because here we had two coupled equations. There were coupled together. We couldn't solve them instantly. That step of elimination reduced me to-- down here in this corner-- one equation, that I've eliminated the first unknown x from the second equation. So the second equation is $0x$ plus $1y$ equal right hand side. And I solve it immediately. OK. So how did I get to that 1 by 1 problem, with these guys removed?

Well-- yeah. I'll just write-- can I write here the-- my parallel way to think of it? 2 by 2 is pretty small, I admit. OK. So I start with 2, 3, 4, 7. I want to split it into-- I want to get the first row and column in one piece. Something goes there. And the other piece is something there. OK. That's what elimination has done. It's taken the original matrix. It's split-- these are both rank 1.

So let's just-- first of all, you could tell me what goes in that blank space in the first rank 1 matrix. So what-- can I say this in words? The first stage of elimination pulls off from A -- so A is some big matrix. It pulls off from A . It takes account of the first column and row. So it writes A as-- here we go-- as a first column, say column 1, row 1, plus the easy part. The easy part will be a matrix with all zeros there-- all zeros there.

And here I have a 2. Can I call it a 2? This is my way now to think about what elimination is really doing. It's starting with an n by n matrix. It's pulling off a rank 1 matrix, which gets that column and that row correct. And it gets whatever it has in here. And then the rest of what's in

there is A_2 . Do you see that we've done that here?

The first step got the first row and column correct. And if it's rank 1, what number goes there?

AUDIENCE: 6.

PROFESSOR: 6. 6 for there. And then this is the rest. This is what we have, one size smaller to work on. And it looks like it was 7. 6 has been used. So it's a 1. That's really-- I want to think of this rank 1 matrix as the first column of I times the first row of u . And then this guy is the second column of I times the second row of u . OK. I haven't presented as proof in a class before. And for 2 by 2, it's looking like overkill to me.

I mean, why? You don't have to do all that deep thinking to get the pieces. But my idea is that it gives the breakdown. And this, of course, is, by our column times row rule, that's LU. So we're starting with A , and we're breaking it up into LU, where lu -- the first piece of lu -- is the first column times row. And then the next pieces are the rest of the matrix. And those get broken down-- the next stage of elimination would break-- if I had a 3 by 3, this stage peeled off the first column and row-- then the next stage would peel off the second-- the new second column and row.

And the third stage would have the third column and row-- just the last pivot does this make any sense to you? You could email me and say it's not that great. But I think it's-- to see that the final result of elimination is I times u is-- there's a little magic in seeing what you're doing. And I think this is a way to see what you're doing-- that you're peeling off a first part to leave a second part like that.

Then the second part, you would peel off the second column times the second row, maybe divide by the pivot to make it correct. And that would put something in the rest of the box. And then A_3 would be the rest of that box. OK. I'm stopping here. I'm glad you let me do 2 by 2, since I see that 3 by 3 would have ruined the day. Yeah. OK. A question, or let me pause for a minute. So I've talked about these factorizations.

This one we won't see again. This one we will see, big time. And this one we will. And this one we will. Yeah. 2, 3 and 5 are the ones that we're really going to see a lot of. Questions or thoughts or-- OK. I guess I want to tell you now to complete today's-- moving forward in this subject-- the fundamental theorem of linear algebra-- the fundamental theorem of linear algebra. OK. Ready for that? Or you may have seen it already, because it's like the highlight of

this subject-- of the basic ideas in this subject.

Right. And then maybe I can-- after I tell you that theorem-- people around the world send me homework problems to do. Now, you would think any sensible professor would never do those problems. He would say, it's your problem. But I get carried away and I solve them sometimes. So one came from India last week, and it involved the fundamental theorem of linear algebra.

Whoever teaching it there really was on the ball. And well, I'll tell you that problem after the fundamental theorem. OK. Fundamental theorem. It's about four sub spaces. So I invented the name four fundamental sub spaces. So can I list the four sub places? Fundamental sub spaces. Well, we know one of them already. The column space-- so, for a matrix.

We are given a matrix A -- that's m by n of rank r . That's our normal starting point. So what are those four sub spaces, and how are they related? And what's their dimension? And what-- those are key facts. OK. We already know the column space-- column space of a matrix. And actually, we already know the row space of a matrix.

And we have the notation for that, column space of A transpose. And what is the dimension-- so that was the key point in the first lecture. Anybody who missed the first lecture, should go back to the notes of 1.1 for the thinking that goes into the dimension equals what? Which of those three numbers do I want to-- is the dimension of the column space? R . And what can I say about r , right away, compared to m and n ?

AUDIENCE: [INAUDIBLE].

PROFESSOR: Yeah. Less or equal. r couldn't-- I couldn't have more independent columns than I have columns. So I've got n columns. So r of them are independent. So r is somewhere, less or equal-- hopefully equal to n . What about the dimension of the row space? How many independent rows has the matrix got?

AUDIENCE: R .

PROFESSOR: R . Thank you. That's the great fact with a new proof last time in section 1.1-- that those have the same dimension, same dimension. Which is-- you think, oh, OK. You look at a simple example. It's true. But if you're given a matrix that's 50 by 100, really the fact that those 100 columns have the same number of independent ones as those 50 row-- that's like great. OK. Now the other spaces are the null space of the matrix, N of A .

And just to make everything naturally symmetric, the null space of A transpose. Those are the last two. Those are the four fundamental sub spaces, which you've seen. And they're even on the cover of the linear algebra textbook. OK. So what's the null space?

AUDIENCE: It's the set of [INAUDIBLE].

PROFESSOR: It's the set of solutions to Ax equals 0 , right. Null space is all solutions to Ax equal 0 . So the null space has vector-- these vectors in it-- the x 's. The null space isn't taken from the matrix. The row space and the column space-- those numbers are sitting in the matrix. The null space, and the null space of A transpose, are solutions to-- the word null is reflecting the fact that that's a 0 . And that's what makes it a space.

Now can you just-- let me just ask you to think again. What's implied when I say-- when I use the word space-- a space of vectors?

AUDIENCE: Closed under addition--

PROFESSOR: I can add-- yeah. So I can do the most important operations of linear algebra in that space. I can add two vectors. Here, let me just add them. So here I'll have a vector x , and let me say another one, a vector y . Then, I do addition. I follow the rules. I see that this can be written as Ax plus y is 0 plus 0 . So what have I learned? I've learned that if x is in the null space, and y is in the null space, then x plus y . So the null space is, as you said, closed, meaning I don't go outside it.

If x is in it, and y is in it, then the sum is in it. And similarly, from Ax equals 0 , I get to A times cx equals 0 . Just multiply by c -- by a number c . So those two facts-- that means I can do linear algebra. I can multiply by numbers. And I can add. In other words, I can take linear combinations. That's what you do with vectors. And the point is, if I do it-- if I take combinations of two null space guys, I'm still in the null space. OK.

So that's the point of the null space. And-- well now-- so now part of the fundamental theorem is to figure out how many independent vectors are in the null space. How many solutions-- independent solutions-- does that system of equations have? So that would be the dimension. And I have to ask you what it is. Let me draw a picture, while you're thinking about those spaces.

It's fantastic to have these beautiful clean boards. OK. So here's my picture of the row space. Row-- that's the column space of A transpose. And here's my picture of the null space, N of A .

And that's the solutions to $Ax = 0$. And why have I put these two together, and these two together? So-- and the other pair will be the column space, C of A . and the null space of A transpose.

So there are the four spaces. Their relationship is the fundamental theorem of linear algebra. So first of all, what-- so I have an m by n matrix. So that tells me that my row-- a typical row has n components, right? I look at an m by n matrix. Let's do a 2 by 3 matrix. So if I look at the row space, this is m . And this is n . So I see three-- the rows have length three. And of course, they multiply the x 's, which also have the length three, x_1, x_2, x_3 .

That's why these are together, because they're both in n dimensional space. Then why are these together? Because the columns are in two dimensional space, for this example. And the null space of A transpose would be just two components, like y_1 and y_2 , to give 0s. So do you see that this is R -- these guys are in R^n ? So that's the first-- get things straight. Two spaces in R^n , two spaces in R^m . Now, what am I going to ask about these spaces?

I guess I've already started asking and didn't wait for an answer. Their dimension. So this has dimension R . And what's the dimension-- how many-- this is really such a key fact. If I have m equations, $Ax = 0$. And if R of those equations are independent, how many solutions? So the dimension of this space is going to tell me how many solutions to $Ax = 0$ in two m equations, but really only are genuine independent equations in this system $Ax = 0$. How many?

So can I ask the question again? And I want the answer in terms of m , and n , and R . So I have-- I really have R equations. If I look at $Ax = 0$, it looks like m separate equations. But m minus R -- of those-- are just copies or combinations of others. So there are independent equations. So what's-- how many have I got?

AUDIENCE: [INAUDIBLE].

PROFESSOR: And And that's what I'm going to write in here. So-- yeah. So x has n components. And there are real active equations that they have to satisfy. And that leaves n minus R . That's the key point. That's the key point that there are n components of x -- n unknowns-- n unknowns. And there are R constraints-- independent constraints. So those n get-- if I want to satisfy those constraints that knocks out dimension R , and leaves n minus R . So that's the dimension here.

And the beauty of the count is that those two numbers add up to n . Everybody's accounted for. Every vector has a piece in the row space and a piece and the null space. And that's-- those

two pieces give you back the vector. Do you see that? That's just nice, that the numbers come out right. And of course, they come out right here, too. You could say, just transpose the matrix and write it-- write the same thing again. What's the dimension of the column space? Equals? The other column space in the matrix has dimension--

AUDIENCE: R.

PROFESSOR: R. Right. And the row-- this guy is left out of some linear algebra books, as if it doesn't belong. But isn't it clear that without it, everything is only three quarters done? We have to have this guy. And its dimension is

AUDIENCE: M minus R.

PROFESSOR: M minus R. Yeah. That count is just, for A transpose, what this count was for A. Yeah. So we've got those dimensions, R and n minus R, R and m minus R. Yeah. You'll have known this, but we need to see it once again in 2018, before we start using it. Now is that the fundamental theorem? Is that all to it there is? No. There is another piece to the fundamental theorem, which is, sort of you could say, the geometry.

Here I have a sub space. Here I have a subspace of this big n dimensional space. So I visualize those sub spaces as some kind of a plane-- an R dimensional plane- and an n minus R dimensional plane. And I want to see how are those two planes connected. How are those two planes connected? And let me get a piece of the-- blank piece of the board to remember the final step. Right. So we've got dimensions r and n minus r, and then over here r, and m minus r. OK. And this is for the rows. This is for the null space.

So this has the rows in it. This has the null-- the solutions to $Ax = 0$. What is the beautiful geometry-- how do you visualize those two spaces? How do you visualize? Let me take an example. Let A be 1, 2, 4, 2, 4, 8. Sorry about that. That's kind of a-- you see it hoked up example. So this is 2 by 3. So there n is 3. What's in the null space of this matrix?

Can you see a vector that solves $Ax = 0$? And in fact, how many will there be? What's-- yeah, what's r for this matrix? Just tell me all the good stuff. For that example, m is 2, n is 3, and r is--

AUDIENCE: 1.

PROFESSOR: 1. Everybody sees 1 for the rank? The rows are dependent. There's only one independent

row. The columns are dependent. There's only-- every column is a multiple of 1, 2. It's a rank 1 matrix. OK. What about its-- so its row space has dimension--

AUDIENCE: 1.

PROFESSOR: 1. And its null space has dimension--

AUDIENCE: 2.

PROFESSOR: 2 So, cause n minus r will be 2. So I'm looking for a couple of vectors that both give 0. I believe there-- I think I've only got one independent row there. So I should be able to find two different vectors that solve Ax equals 0. So what what's the solution to Ax equals 0?

AUDIENCE: 0 minus 2, 1.

PROFESSOR: 0 minus 2, 1. Yeah, that works. And what's an independent solution?

AUDIENCE: 4, 0, negative 1?

PROFESSOR: 4, 0-- don't throw me off-- 4, 0 and--

AUDIENCE: Minus 1.

PROFESSOR: Minus 1. Yeah. That looks good. That looks good. And then the claim is that every solution would be a combination of those two. And this is how many there are. And now, it's the geometry I'm completing. So we have two minutes left in this lecture. You just have to tell me how-- what's the relation between these guys in the row space, and that guy in the null space? What's the relation between the rows of A , the solutions to Ax equals 0?

Between-- if you see it-- if you saw that vector and that vector-- well, A times x is 0. So what does that tell us? What do we see for the relation between 1, 2, 4, and 0 minus 2, 1.

AUDIENCE: Orthogonal

PROFESSOR: They are orthogonal. Terrific. Yes. Orthogonal I test by the dot product, 0 minus 4, 4, add to 0. Yes. So the-- and that's a completely general fact. When I look at Ax equals 0, it's telling me that x is orthogonal to the rows. Do you see that? Just to put it in again here. If I look at Ax , A has a bunch of rows. x has one column. And I get 0. That's the point of the null space. And that equation is just saying that row 1 is orthogonal, because that's the dot product of row 1 with x .

So here is row 1, row 2, row 3, and row 4 with x . The rows with x -- and I get 0s. So the point is then, these two spaces are at 90 degree angles. That's really a neat picture of the four sub spaces. And these two are for the same reason-- at 90 degree angles-- off in m dimensional space. So this is the fundamental theorem of linear algebra-- to see that the dimensions come out right, and the geometry comes out right. Yeah.

And then, now, next time-- following the notes-- and I have a few more copies of the one hand out. We'll move on quickly next week to eigenvalues and positive definite matrices. Good this is really linear algebra moving on.