GILBERT STRANG: OK, ready for part three of this vision of linear algebra. So the key word in part three is orthogonal, which again means perpendicular. So we have perpendicular vectors. We can imagine those.

We have something called orthogonal matrices. That's when-- I've got one here. An orthogonal matrix is when we have these columns. I'm always going to use the letter Q for an orthogonal matrix. And I look at its columns, and every column is perpendicular to every other column.

So I don't just have two perpendicular vectors going like this. I have n of them because I'm in n dimensions. And you just imagine xyz axes or xyzw axes, go up to 4D for relativity, go up to 8D for string theory, 8 dimensions. We just have vectors. After all, it's just this row of numbers or a column of numbers.

And we can decide when things are perpendicular by that test. Like say the test for Q1 to be perpendicular to Qn is that row times that column. When I say times, I mean dot product, multiply every pair. Q1 transpose Qn gives that 0 up there. So the columns are perpendicular. And those matrices are the best to compute with. And again, they're called Q.

And one way to, a quick matrix way, because there's always a matrix way to explain something, and you'll see how quick it is here. This business of having columns that are perpendicular to each other, and actually, I'm going to make those lengths of all those column vectors all 1, just to sort of normalize it. Then all that's expressed by, if I multiply Q transpose by Q, I'm taking all those dot products, and I'm getting 1s when it's Q against itself. And I'm getting 0s when it's one 1 Q versus another Q.

And again, just think of three perpendicular axes. Those directions are the Q1, Q2, Q3. OK? So we really want to compute with those. Here's an example. Well, that has just two perpendicular axes. I didn't have space for the third one.

So do you see that those two columns are perpendicular? Again, what does that mean? I take the dot product. Minus 1 times 2, 2. 2 times minus 1, another minus 2. So I'm up to minus 4 at this point. And then 2 times 2 gives a plus 4. So it all washes out to 0.

And why is that 1/3 there? Why is that? That's so that these vectors will have length 1. There

will be unit vectors. Yeah, and how do I figure, the length of a vector, just while we're at it?

I take 1 squared or minus 1 squared gives me 1. 2 squared and 2 squared, I take the dot product with itself. So minus 1 squared, 2 squared, and 2 squared, that adds up to 9. The square root of 9 is the length. I'm just doing Pythagoras here.

There is one side of a triangle. Here is a second side of a triangle. It's a right triangle because that vector is perpendicular to that one. It's in 3D because they have three components. And I didn't write a third direction. And their length one vectors because just that's how when I compute the length and remember about the 1/3, which is put in there to give a length 1. So OK.

So these matrices are with Q transpose times Q equal I. That again, that's the matrix shorthand for all I've just said. And those matrices are the best because they don't change the length of anything. You don't have blow up. You don't have going to 0.

You can multiply together 1,000 matrices, and you'll still have another orthogonal matrix. Yes, a little family of beautiful matrices. OK, and very, very useful. OK, and there was a good example.

Oh, I think the way I got that example, I just added a third row. The third column, sorry. The third column. So 2 squared plus 2 squared plus minus 1 squared. That adds up to 9. When I take the square root, I get 3. So that has length 3. I divided by 3. So it would have length 1. We always want to see 1s, like we do there.

And if I-- here's a simple fact. But great. Then if I have two of these matrices or 50 of these matrices, I could multiply them together. And I would still have length of 1. I'd still have orthogonal matrices. 1 times 1 times 1 forever is 1.

OK, so there's probably something hiding here. Oh, yeah. Oh, yeah, to understand why these matrices are important, this one, this line is telling me that, if I have a vector x, and I multiply by Q, it doesn't change the length. This is a symbol for length squared.

And that's equal to the original length squared. Length it is preserved by these Qs. Everything is preserved. You're multiplying effectively by the matrix versions of 1 and minus 1. And a rotation is a very significant very valuable orthogonal matrix, which just has cosines and signs.

And everybody's remembering that cosine squared plus sine squared is 1 from trig. So that's

an orthogonal matrix. Oh, it's also orthogonal because the dot product between that one and that one, you're OK for the dot product. That product gives me minus sine cosine, plus sine cosine, 0. So the column 1 is orthogonal to column 2. That's good. OK. These lambdas that you see here are something called eigenvalues. That's not allowed until the next lecture.

OK, all right, now, here's something. Here's a computing thing. If we have a bunch of columns, not orthogonal, not length 1, then, often, we would like to convert them to, so we call those, A1 to AN. Nothing special about those columns. We would like to convert them to orthogonal columns because they're the beautiful ones, Q1 up to Qn.

And two guys called Graham and Schmidt figured out a way to do that. And a century later, we're still using their idea. Well, I don't know whose idea it was actually. I think Graham had the idea. And I'm not really sure what Schmidt, how he got into it. Well, he may have repeated the idea.

So OK, so I won't go all the details. But here's what the point is the point is, if I have a bunch of columns that are independent, they go in different directions, but they're not 90 degree directions. Then I can convert it to a 90 degree one to perpendicular axes with a matrix R, happens to be triangular, that did the moving around, did take that combinations.

So A equal QR is one of the fundamental steps of linear algebra and computational linear algebra. Very, very often, we're given a matrix A. We want a nice matrix Q, so we do this Graham Schmidt step to make the columns orthogonal. And oh, here's a first step of Graham Schmidt. But you'll need practice to see all the steps. Maybe not.

OK, so here, what's the advantage of perpendicular vectors? Suppose I have a triangle. And one side is perpendicular to the second side. How does that help?

Well, that's a right triangle then. Side A perpendicular to side B. And of course, Pythagoras, now we're really going back, Pythagoras said, a squared plus b squared is c squared. So we have beautiful formulas when things are perpendicular.

If the angles are not 90 degrees when the cosine of 90 degrees is 1 or maybe the sine of 90 degrees is 1, yeah, sine of 90 degrees is 1. For those perfect angles, 0 and 90 degrees, we can do everything. And here is a place that Q fits. This is like the first big application of linear algebra. So let me just say what it is.

And it uses these cubes. So what's the application that's called least squares? And you start

with equations, Ax equal b. You always think of that as a matrix times the unknown vector, being known, right hand side b Ax equal b.

So suppose we have too many equations. That often happens. If you take too many measurements, you want to get an exact x. So you do more and more measurements to b. You're pasting more and more conditions on x. And you're not going to find an exact x because you've got too many equations. m is bigger than n.

We might have 2,000 measurements, say, from medical things or from satellites. And we might have only two unknowns, fitting a straight line with only two variables. So how am I going to solve 2,000 equations with two unknowns Well, I'm not.

But I look for the best solution. How close can I come? And that's what least squares is about. You get Ax as close as possible to b. And probably, this will show how the-- yeah. Yeah, here's the right equation.

When you-- here's my message. When you can't solve Ax equal b, multiply both sides by A transpose. Then you can solve this equation. That's the right equation.

So I put a little hat on that x to show that it doesn't solve the original equation, Ax equal b, but it comes the closest. It's the closest solution I could find. And it's discovered by multiplying both sides by this A transpose matrix.

So A transpose A is a terrifically important matrix. It's a square matrix. See, A didn't have to be square. I could have lots of measurements there, many, many equations, long, thin matrix for A. But A transpose A always comes out square and also symmetric. And it's just a great matrix for theory. And this QR business makes it work in practice.

Let me see if there's more. So this is, oh, yeah. This is the geometry. So I start with a matrix A. It's only got a few columns, maybe even only two columns. So its column space is just a plane, not the whole space.

But my right hand side b is somewhere else in whole space. You see this situation. I can only solve Ax equal b when b is a combination of the columns. And here, it's not. The measurements weren't perfect. I'm off somewhere.

So how do you deal with that? Geometry tells you. You can't deal with b. You can't solve Ax equal b. So you drop a perpendicular. You find the closest point, the projection that in the

space where you can solve. So then you solve Ax equal p. That's what least squares is all about, fitting the best straight line, the best parabola, whatever, is all linear algebra of perpendicular things and orthogonal matrices.

OK, I think that's what I can say about orthogonal. Well, it'll come in again. Orthogonal matrices, perpendicular columns is so beautiful, but next is coming eigenvectors. And that's another chapter. So I'll stop here. Good. Thanks.