[MUSIC PLAYING]

GILBERT STRANG: So, I'm Gilbert Strang. And this is about my new course 18.065 and the new textbook Linear Algebra and Learning from Data, and what's in those subjects. So there are really two essential topics and two supplementary, but all very important subjects. So if I tell you about those four parts of mathematics that are in the course, that will give you an idea if you're interested to follow through. So the first big subject is linear algebra. That subject has just surged, exploded in importance in practice.

What I want to focus on is some of the best matrices, say, symmetric matrices, orthogonal matrices, and their relation. Those are the stars of linear algebra. And the key step is to factor a matrix into maybe symmetric times orthogonal matrix, maybe orthogonal times diagonal times orthogonal matrix-- that's a very important factorization called the singular value decomposition. That doesn't get into a lot of linear algebra courses, but it's so critical.

So can I speak now about the second important topic, which is deep learning? So what is deep learning? The job is to create a function. Your inputs are, like for driverless cars, the input would be an image that's a telephone pole or a pedestrian. And the system has to learn to recognize which it is. Or the inputs from handwriting on addresses would be a zip code. So the system has to learn how to recognize 0, 1, 2, 3 up to 9 from handwriting of all kinds.

Or another one is speech, like what Siri has to do. So my speech has to get input and interpreted and output by the process of deep learning. So it involves creating a learning function. The function takes the input, the data, and produces the output, the meaning of that data. And so what's the function like? That's what mathematics is about, functions.

So it involves matrix multiplication. Part of the function is multiplying vectors by matrices. So that's a bunch of steps. But if there was only that, if it was all linear algebra, the thing would fail and has failed. What makes it work now so much that companies are investing enormously in the technology is that there is now a nonlinear function, a very simple one in the middle between every pair of matrices. And that nonlinear function, I can even tell you what it is.

It's a function f, let's call it f, f of x is equal to x if x is positive. And f of x is 0 if x is negative. So you can imagine it's graph. It's a flat line where it's negative. And then it's a 45 degree slope where it's positive. So putting that nonlinear function in between the matrix multiplications is

the way to construct successful learning function. But you have to find those matrices.

I mentioned two supporting subjects. The first is-- optimization would be the word. We have to find the entries in those matrices that go into the learning function. That's a crucial step. So this is a problem of minimizing the error with all those matrix entries as variables. So this is multivariable calculus, like 100,000. 500,000 variables, it's just unthinkable in a basic calculus course, but it's happening in a company that's working with deep learning. And so that's the giant calculation of deep learning. That's what keeps GPU's going for a week. But it gives amazing results that could never have been achieved in the past.

So then the other key subject is statistics. And the basic ideas of statistics play a role here, because when you're multiplying a whole sequence of matrices in deep learning, it's very possible for the numbers to grow out of sight exponentially or to drop to zero. And both of those are bad news for the learning function. So you need to keep the mean and variance at the right spot to keep those numbers in the learning function, those matrices in a good range.

So this course won't be a statistics course, but it will use statistics as deep learning does. So those are the four subjects. Linear algebra and deep learning, two big ones. Optimization and statistics, essential also.

So I hope you'll enjoy the videos and enjoy the textbook. And go to the OpenCourseWare site ocw.mit.edu for the full picture. Beyond the videos, there are exercises, problems, discussion, lots more toward making a complete presentation, which I hope you like.