PROFESSOR STRANG:

Ready for the least squares lecture, lecture 11? Homework is just being posted on the web. It'll be due, it's really to help you practice, get some experience on these sections for the first exam. That's Tuesday evening. So eight days away. So the homework will be due the day after. And actually, we'll try to move the review session to Monday next week so you can ask me any questions about the homework or any review material. So that's all a week away and this week we get two great examples. Least squares is one that comes today.

But could I first, because I keep learning more-- And I've got your MATLAB homeworks to return. I keep sort of learning a little more from your MATLAB results and I think because we spoke about it, it would be worth speaking just a little more. So I'm going to take ten minutes about this convection-diffusion equation in which I put in a coefficient d, a diffusivity, just to help get the units right. So this is your example. And it had d=1 of course. Well first I realized that later in the book I completely forgot that I discussed this problem. About page 509, I think. I discussed it a little bit. And just because it's worth, since we invested a little time, the little bit more will pay off. So first of all, the point is here we have convection competing with diffusion.

And always there's some non-dimensional number. Here it's called the Peclet number. Actually, there's an accent on one of those e's, PÃ©clet number. Which measures the ratio, the importance of convection relative to diffusion. So it's V times a length scale in the problem, divided by d. So then that has the same units as that, if the result is dimensionless. Maybe you know the Reynolds number. This is very like the Reynolds number, which also measures, in Navier-Stokes equation, the importance of convection, advection, and diffusion. There, in that equation, the velocity V, that's a non-linear equation, Navier-Stokes, it's tremendously important and many codes to solve it, lots of discussion, theory still not complete. In that problem, the V is u. It's non-linear and the term there that we took as a constant, as a given constant V, it's the same as u. So in the Reynolds number, this would be u, a typical velocity u, times a typical length scale, which would be like one in our zero to one problem, divided by d or mu or nu, whatever number we use. So it's like the Reynolds number.

And then it's turned out for this problem that people also use a number that gets called the cell Peclet number where the length is taken to be half the cell size, delta x over two. Let me call that number P. So that's P. And what's my point? This equation's important enough to sort of see a little more about it than just the numbers that come out. So the MATLAB homework, which you did really well, set up finite differences for this. Right? And found the eigenvalues and solutions. It's the eigenvalues I want to say a little more about. Because you set up a matrix K over delta x squared and V times the centered difference over delta x. And I guess I call that whole combination L, and asked you about the eigenvalues of L. And you printed them out correctly.

But there's more there than I think we have understood. And I want to make some more comments about that. Because it's quite important. And the comments are clearest if I just reduce to a n equal two. So that matrix, well the off-diagonal part of that matrix had some number b and some number c. Actually we could figure out what was the b in this. This produced a minus one, two, minus one, right? So part of the b was the minus one over delta x squared. And then from this was a plus V and a one over, well it's a centered difference so I should divide by 2 delta x. Is that right? Is that what a typical off-diagonal thing in the matrix that you displayed? That's what's coming from the off-diagonal of K. And this is what's coming from the centered difference C. And then what would this c thing be? Well the c is below the diagonal so it's also that minus one over delta x squared. But now this is a difference, so it's going to be a minus, right? I think those would have been your entries for b and c.

So can we just think first, what are the eigenvalues of that matrix? It's a two by two, simple problem. The trace is zero plus zero. So that the eigenvalues will be a plus minus pair because they have to add to zero. And I think that's the plus minus pair you get. Let's just check. What's our other check? They will add to zero, the plus the square root and minus the square root. And the product of the two eigenvalues, lambda_1 times lambda_2, will be, we have one of them is plus, one with a minus, so it'd be minus bc. And that's correctly the determinant. So it's good. These are the correct eigenvalues.

Now let me ask you about the signs of b and c. If b and c have the same signs, like maybe even equal, one, one, what are the eigenvalues? So in that symmetric case if b and c are equal the eigenvalues are? Right here. If b and c are equal, say equal to one, the eigenvalues are plus and minus one. But what if the signs are opposite? Everything changes. What if b is one and c is minus one? That matrix would then be a 90 degree rotation. It would be anti-

symmetric if b was one and c was minus one. Our formula is still correct, but what does it give us? If b is one and c is minus one what have I got here? I've got i. So the eigenvalues change from plus and minus one in the symmetric case to plus and minus i in the anti-symmetric case. And I think that's what you guys saw at a certain level of V. I hope you did because that was the point about eigenvalues.

Now you may say, what about the diagonal? Well I claim diagonal is very simple. What's the diagonal? Now I'm going to allow myself a diagonal and I'm just going to change-- What happens if I have a and a? Same entry on the diagonal. What are the eigenvalues now? This is just like, a great chance to do some basic eigenvalue stuff. What are the eigenvalues of that matrix? Well I've added a times the identity. I've just shifted that matrix by a. So the eigenvalues all shift by a. So the eigenvalues are now a plus and minus. So no big deal. So you say that the a is actually not important, not the key to this question of are they real or do they go complex. So the eigenvalues of this are real when b and c have the same sign. If b and c have the same sign, I have a square root, no problem. When b and c have opposite sign, what do I get? When b and c have opposite sign, I'm taking the square root of a negative number and I've gone complex. Do you see that the change from real eigenvalues, which gives a nice curve, to complex eigenvalues, which gives a very bumpy curve for the solution, just happens when like, for example, b-- Is it b that's going to go to zero maybe? And then beyond that?

Well this sign is for sure negative, right? So c is staying negative. And originally for a little delta x, b is also negative. What's happening here? I think that the transition that you I hope observed comes when b hits zero. When the combination of V and delta x is such that at b=0 we switch from real eigenvalues to complex eigenvalues. And when is b=0? That's when this negative guy off the diagonal just exactly cancels this one. So b is zero when what? So if this equals this, one over delta x squared is equal to V over two delta x. Let me multiply both sides by delta x squared so that I have a nice one there, multiplying by delta x squared will put a delta x up here.

And what have we discovered? This is why I wanted you to see it. That the transition comes when the Peclet number is one. So that Peclet number, that cell Peclet number is exactly the point that we observed of transition from real eigenvalues to complex eigenvalues. And that's the transition. So it's that combination, this is the Peclet number, cell Peclet number, it's that combination, P_cell maybe. We've done the computations and now we gradually get back to

the meaning. And I just wanted to take this step back to the meaning to see when do those numbers start going complex. You may have noticed or you may not have noticed that it'll happen when one of those, when that upper diagonal changes sign.

Now you could say, okay that's the eigenvalues. What's the consequences for the shape of the solution? Well, I haven't figured all that out. I'd be happy to have some more thoughts about that. But what you noticed, I think, in the computations is if V got too big so that that P was bigger than one, if V got too big, so convection was dominating and our delta x was not small enough to deal with it, you should have seen the points on the discrete values were oscillating instead of a proper smooth-- I mean, the proper, with a large V, the correct solution, I think, is practically nothing for here and then it goes, this is a really large V, take V to a thousand or something. It climbs up like mad. Here's the halfway point where the load is. And then it goes along here and then it climbs down like mad to satisfy the boundary condition. I didn't know that that's what would happen for large V. What I'm saying is, and undoubtedly it could be understood physically, so I guess what I'm saying is there's just more good stuff in any computation than purely the numbers. And this is part of the good stuff in that example.

I hope you liked that. Because I mean, here you did the work but then, to understand it is frankly still under way. More thinking to do. That's back to least squares. Here's today's lecture. So remember where we started last time. Au=b. Last time I wrote f. I regret it terribly. I can't fix it. But it's b. I want b there to be the right-hand side. And I jumped to the equation that determines the best u. There's no exact u because we've got too many equations. You remember the set-up, we have too many equations. There's noise in the measurements and we can't get the error down to zero. There's some error. And the best u was given by that equation and we want to say why. And understand it from two or three ways. Calculus, geometry, everything.

Can I first, because I love my little framework here, fit it in because it's quite important, this example and then others fit in. So u is our unknown as always. Then the matrix A in the problem produces an Au. Now two things to notice about e, which, that's the same letter I used for elongation, here it's standing for error. Two things to notice. One is that the source term, which is b, comes in at this point of the framework. When we had external forces on springs and on masses it came in at this point. We had an f there. So that's why I'd like to keep those two separate. The b's are like voltage sources, they come in here. The f's are will be like current sources, they'll come in there. Actually it's beautiful.

One more thing to notice. A is coming with a minus sign. In mechanics, in masses and springs, we had e=Au. Here it's natural to work with this, the error or the residual b-Au. And that minus sign is natural in physics and in electrical engineering and hydraulics, you know, flow-- Where's that minus sign coming from in flow? Well, flow goes from the higher point to the lower. Higher voltage to the lower voltage. And that usually produces that minus sign. No big deal, of course. So that step is fine with the framework.

What do we expect in that middle step? So what's our name for the matrix that goes there? Everybody's gotta know this framework. C, right? Only I've been taking unweighted least squares. So for unweighted least squares, C will be the identity. And C doesn't show in our equations. So C is the identity when there are no weights, when all the equations are equally reliable. And that's pretty common, of course. But not always. And we'll think, okay, there is a weight e. So w, which is Ce, is weighted errors, you could say. So the letter w comes up appropriately again. Weighted errors. And then what's the good weighting? May I stay with C equal the identity for the moment? Unweighted least squares, because that's by far the most common. And then w and e are the same. C is the identity. And finally, there's the last step in our framework where we always expect to see A transpose. And we do. And we have to say why.

So that's where I left it last time. That this was the picture. This is the equation. If I had a matrix C, it would go there and there. Right? Because I'd have b-Au and then I'd apply C before A transpose. So C would slip in there before A transpose on both sides. So that would, with the C's there, that would be the weighted least squares equation. You see that it would be A transpose C A instead of A transpose A, but still the main facts are there.

So where does the equation come from? So one source, one way to get the equation is from calculus. From minimizing, from minimizing. Set a derivative to zero, calculus. And what's the quantity we're minimizing? We're minimizing that squared error because this is least squares. We're minimizing this e transpose e, the length of e squared. The sum of the squares of the errors. Which is (b-Au) transpose (b-Au). Again I could say where to slip in the C matrix. If there was one, it would go in there. C would go in there, C would go in there. There'd be a C in the equation. But let's keep C to be the identity. So I minimized. It's a quadratic. It's got u's times u's, so second degree. And what's the coefficient in that second degree part? Well, the second degree part is coming from $(Au)^\wedge$ transpose Au. Right? This times this is going to be linear. This times this is going to be linear. That times that is just going to be a constant, its

derivative is zero. But this times this is altogether, that times that is the u transpose A transpose Au. Right? So that's the quadratic part. And my only point is it's like our old stiffness matrix. We're seeing the matrix in here is A transpose A. In other words, when I do calculus and maybe I'd prefer to see something than just compute away, take derivatives mechanically. So I'm going to leave that which is done in the text, finding the derivative, setting to zero. And what does it give? It gives us our equation. So that equation will come when I set the derivatives of this thing to zero. So that's one totally okay approach.

But I like to see a picture with it. I hope that's alright. To take the second approach is to see why A transpose w equal zero. Why is that? What's going on in that key step? This is always the key step. This is like the set-up step. This is the weighting step with constants coming in. And here's the key step. Let's see that. So my picture. Let me draw that picture again. And my example was in three dimensions, so m=3. I've got three equations. The matrix A, oh I'm afraid I don't remember what it was, but I think it was something like 1, 1, 1; 0, 1, 3, was that maybe it? Just to connect to last time. And what I'm now calling b was the vector [1, 2, 3] was it? Or was it not? It was [0, 1, 2] maybe? That's right? And what was the point? If I draw the vector b it goes there somewhere. If I draw the first column of A, it goes here somewhere. If I draw the second column of A, it goes there somewhere. And if I draw all combinations of these columns, all combinations of that vector and that vector, what do I get? I get a plane. I get a plane. There it is. That's the plane. That's the plane. This is from column one. Here's column two. This plane is the column plane or column space. It's the column space of A because it comes from the columns of A.

Now what's the point about this plane? The point is that if b is on the plane then I'm golden. If b is on the plane then b is a combination of the columns, that's what the plane is, and I have a solution to Au=b. So b on a plane, b on the plane means Au=b is solvable. And it could happen, of course. Like perfect measurements. But we can't expect it. When we have three measurements or 100 measurements or 10,000 measurements we can't expect perfection. So usually b will be off the plane. Now what? What happens when b is off the plane? Let me just complete that picture. And you know what's coming. If we're going to get-- Au or Au hat is going to be on the plane so I'm looking for the best u hat.

Can I just erase this to make space for what you know I'm going to draw? Here are these little columns, let me put them there. What am I going to draw? The projection. The projection. I'm going to draw, what's the projection? The projection is the nearest point that is in the plane to

the b that's not in the plane. So here's the projection p. I drop down this thing. There's the projection p, little p. That's the projection of b onto the plane. I think your mind says yeah, that's the right choice. And do you want to tell me what this? That is the part that we can't deal with. The part we can't improve. We've made it as small as we could and it's e. That's the error e and this p is the best guy that is in the plane. Do you see that this is the picture. You get an actual picture of what's going on. You're splitting b, the measurements, into the part you can deal with, the projection, the Au hat that is in the column space. It is a combination of the columns. Those points do lie on a line if I'm doing straight line fitting. And the part that you can't deal with, the e, the difference, b-Au, which is not in the plane.

And now I'm still looking for the equations. Right? I've just named some stuff. But I haven't got an equation for that projection. So what's the key fact? What's the key fact in this picture that's going to lead me to an equation for p and e and u hat and everything? The key fact is that that dotted line is perpendicular, perpendicular to the plane. If I'm looking for the closest point, everybody knows project, that's what projection involves. Go perpendicular. This is a right angle. That e is perpendicular to the whole plane. Not only perpendicular to p, it's perpendicular to everybody in that plane. Right? I'm dropping the perpendicular to the plane. Do you accept that? Because if you do, we're through. We just write down the equations for perpendicular and we've got what we want from the picture instead of from a calculation.

So what's the idea? So e is perpendicular to the first column. So b in the plane, we would be golden. Let's suppose we're not in the plane. So now we have this 90 degree angle, this perpendicular projection. And it tells me that the first column-- oh I better name the columns. Can I just call this column a_1? That first column is a_1 and the second column is a_2. So those two columns, whatever they are, are the guys whose combinations give us the plane. And it's the plane that we're projecting onto. It's the plane of all combinations that comes up here. So what's this 90 degree angle? It says that a_1 is perpendicular to p, right? Sorry! Say that right for me. The first equation says that a_1 and what are perpendicular? e, thank you, e. So the first equation says that a_1 transpose e is zero. And the second equation says that a_2 transpose e is zero. Those are my two equations.

I have to convert those now into matrix language because I've done them two separate-- vector, I mean vector language, and I want to get into matrix language. But it's easy to do. Here I have, look, if I have two equations, let's get a matrix here. What's it saying? a_1 transpose and a_2 transpose, what are those? They're the rows of A transpose. So the matrix

way to say that is A transpose e equal zero. In other words, this is saying both at once, right? The first row of A transpose times e gives zero, the second row of A transpose times e is zero. So it's A transpose e equal zero which is what we wanted in this case where w and e are the same. Because C is the identity. And let's just go one step further and see. That's A transpose (b-Au hat) is zero. Remember this zero stands for [0, 0], right? I wanted to put the two equations together. So I've got two components on the right-hand side. And then I just plugged in what b is. And now everybody sees it, right? Everybody sees that we've got the picture, this 90 degree angle was the key to these equations. Because if I put A transpose A u hat onto the other side, I've got exactly the normal equations that I wanted.

We're taking the time to see the picture and the form of the equations. Then I can plug in the numbers, but the thinking is where the equations come from. We're there. Now what to do next? Now we've understood where the equations come from. I didn't go through the steps of taking the derivatives, but that would work. Or this picture. I love this picture. Let me stay with that a little bit longer. What is u hat? Can I just go over here to say, okay what have we got here? We started with Au=b and then we got the projection was A u hat. But now what is u hat? I'm just going to assemble things here. u hat, we figured out by the 90 degree angle, comes from this equation, which is that equation, which is A transpose A u hat equal A transpose b, the central equation. That's the central equation.

Now plug in u hat here so I get a formula for the projection. While we're doing all this stuff we might just as well put those two pieces together and have a formula for the projection. So it's A times u hat-- I hope you like this formula. It's kind of goofy-looking but you'll remember it. What is u hat? The whole point is that this matrix is good. It's square, it's symmetric, it's invertible, we'll have another word about that. And now I'll invert it. Times A transpose b. That's the goofy formula that I wanted you to see. The projection of vector b onto these columns of A comes from applying this matrix, sometimes I call it the matrix of four A's. Now it's worth looking at that matrix. Often I'll call that matrix capital P. It's the projection matrix. You give me any vector b, I multiply it by this matrix and I get the projection. It's just worth seeing what this matrix P, these four A's, what are projection matrices like.

Now first of all, when I have an inverse of a product, any reasonable person would say okay, split that into A inverse times A transpose inverse and simplify the whole thing. And what will happen? It's not going to be legal, but let's just pretend. If I split this into A inverse times A transpose inverse and simplify, what do I get for P? Do you see it? I'll get A and if I try to split

this into that, what do I have here? I've got the identity. That's the identity. That's the identity. The result is the identity. That doesn't look good, right? p is not the same as b. This matrix cannot be split into these two pieces. A is rectangular, that's its problem. If A was square, oh yeah, think about the case when A is square. Suppose m equals n. That case'll be included here. If m equals n and my matrix is square and invertible and golden then all this works. The projection is the identity matrix. And what's with my picture? What's my picture look like in the case where A is a square matrix? Give it another column. Fit this thing by a quadratic. So if I was fitting instead of by a straight line, by a quadratic, it turns out I'd have zero squared, one squared and three squared in that column. I'd have a three by three matrix. It comes out to be invertible.

Now what's going on? Now what's my problem Au=b? Now suddenly m is still three, but now n is three. b is? And what happened to the plane? b is in there. And now what's there? It's now the combinations of what? Why did that plane come in? That was the combinations of two columns. But now I've got three. The combinations of three columns, those three columns of an invertible matrix is what? Are you with me? If I have a three by three invertible matrix, these three columns independent, pointing off different directions, not in a plane, then when I take the combinations I get? I get R^3. I get the whole space. I get everybody. Every vector including this b and any other b you want to suggest will be a combination of these three guys. So what's my picture here? My picture is that plane grew to be the whole space. So what's the projection of b onto the whole space? b itself. And what's the error? Zero. Good. So that's the nice case. That's the standard case that we've thought about in the past when m equalled n. In that case P is the identity and that'd be all true. But normally it's not. So I want to come back to this P just to mention an important fact about P. And it comes again from the picture. So this is a projection. This is what I'm calling the projection matrix. It's the matrix that does the projection. And there it is. Four A's in a row that multiplies b. Now here's my little question. So linear algebra's full of these different kinds of matrices. Rotations, reflections, symmetric matrices, Markov matrices, so it's just every problem has matrices. Now here we have a projection matrix. Now what I want to know is what happens if I project again? If I take the vector b, any vector b, I project it and then I project again. So project twice and just tell me, you know what will happen. I'm back to this picture. I project b to P and now I project again. Where do I go? Same place, right? Once I'm in the plane the projection stays right where it is. So what does that tell me? That tells me that P squared on b is the same as P on b. If I project twice, no change. It's the same as projecting once. So the projection matrix has the property that P squared is P.

And actually, we should be able to see it if I write out this whole miserable thing twice. So now I'm going to be up to eight A's. Sorry about this, but I promise not to do P cubed. A times (A transpose A) inverse times A transpose, that's one P. I'll write it again. There's the second P. So that's P squared. Do you see anything good there? Do you see in here A transpose A, that combination and that combination there. This cancels that to give the identity. And what am I left with? I'm left with A, the inverse times A transpose, which was exactly P. The algebra is just coming along with the understanding that we know. So that's the projection matrix. So this is the theory of projections in a nutshell, in a nutshell. This is projections onto the column space of A.

Now I have to remind you about one little math point. Not so little, I guess. How could I say little for math? Is A transpose A invertible? We're plowing along as if it is, that's going to be our assumption. But what's the condition for A transpose A to be invertible, which allows all this to work? When is A transpose A invertible? What I'm doing here now is I'm separating the positive definite one, when A transpose A is positive definite, the good normal case when all our equations work, from the semi-definite one where we overlooked the fact that A transpose A-- where somehow the experiment wasn't well set up, we got an A transpose A that is singular. And just to see, when could that happen? Let me just remind you. This is important. Why don't I give it some space.

It's really straightforward. Let me just go through those steps again. If it's not invertible, if some A transpose A u is zero. This is always the risk that we have to check out and be sure we don't have and understand. So if A transpose Au is zero, then that would lead us, I could multiply both sides by u transpose. u transpose zero, right? Safe. Multiply whatever that u might be, multiply both sides by u transpose. But what is u transpose zero? Zero, nothing there. Now how do I understand this guy? Well you remember the key. Everybody remembers the key? You look at that thing and you say hey, if I put in parentheses in the right place that's the length of Au squared. So that's the small trick that this multiplying by u transpose and then seeing what you've got that we've done and you should know it. And now if the length squared is zero, what does that tell me about Au? If I have a vector here whose length is zero, that vector must be? Zero. Zero vector's the only one for which the sum of the squares will give zero. And if Au is zero I could multiply both sides by A transpose and complete the loop. Actually I thought of that when I was swimming this morning, that line.

Just to see once again when, it's sort of interesting then. A transpose Au equal zero which is

the bad thing we hope we don't deal with. And when does it happen? It happens when Au is zero. So our assumption always has to be this, that there aren't any u's, except the zero vector of course, that's always going to happen, but we always have to assume that Au is never zero. So we have to avoid this. So to avoid that assume A has -- this is the key word -- independent columns. Since this is a combination of the columns, independent columns means what? It means that the only combination of the columns to give zero is the zero combination. So did I have independent columns over here? I sure did. That column and that column were off in different directions, they were independent. And that's why I knew we were fine. A transpose A was zero. I would have to really struggle to find a, well I'd have to think a bit to find an example where we run into trouble. These squares, well I certainly could in many applications, but the straightforward applications of fitting a straight line, A is going to be a column vector of ones and a column vector of times and those are different directions and no problem. So that's A transpose A.

What else to do with this topic? Because there's a whole world of estimation. I mean, statistics is looking over our shoulder I guess. Really, we should realize that a statistician is, say, yeah, I know that but, and then going on. And what is that guy, what more does he have to say? So you've got the central ideas. I guess the statistician comes in in this, that's the statistical constant now. And what do statisticians compute? They say you've got errors, right? And of course, in any particular case we don't know what that error is, otherwise we could take it out and we'd get exact solutions. We don't know what the error is. What is reasonable to know about errors? We're doing a little statistics here. Somehow that error, that particular error of the experiment that we happen to run, and if we ran it again we'd get a different error, those errors come out of some sort of error population. Like dark matter or something. Just like, a bunch of errors are out there, noise. And what could we reasonably assume that we know about the noise? We could assume that its average is zero, mean zero. So statisticians always, that just resets the meter. Right? If you had a meter or a clock that was always three minutes ahead (like this one) you would reset it. And we'll do that one day. So you'd reset to get the average zero.

But that doesn't mean every error is zero, right? That just means the average error is zero. So what's the other number? What's the other number that statisticians live on? It's the deviation or its square, which is called the variance. Right. Variance. So that's the thing that you could assume that the errors have mean zero and have some variance. You could suppose that you knew something about the variance. You don't know the individual errors, but you know

whether the errors are like, are very small or close to zero or large. So this is a small variance. So one over sigma is sort of that distance. One over sigma. Here, this is a large variance where the magnitude of the error could be much larger from this. So those are the two numbers, mean zero, that leaves us just one number, and the variance, the standard deviation sigma or the variance sigma squared.

One moment on these squares. Let me just say what the weighting matrix would be. And then I can tell you in a moment why. What would the weighting matrix be if our three equations, you know, that came from one measurement and this came from a second measurement and this came from a third measurement. If they came from different meter readers with different variances, suppose, then the right C matrix will be a diagonal matrix, beautiful. And what sits up there, what sits there, what sits there? We don't have spring constants anymore. We have statistics constants. And what's the number that goes there? That one is the third guy. So it's associated with the third measurement. It's one over sigma_3 squared. Those are the numbers that go on the diagonal, the inverses of the variances.

And just to see that that makes sense. If that number is unreliable, if it has a large variance then I want to give it little weight, right? If this third meter is very unreliable I'm not going to throw it out entirely, but I know that it's variance is large and therefore I'll weight that equation only a little, with a small weight. Suppose sigma_2, so this guy is one over sigma_2 squared, suppose this is an extremely reliable meter. That measurement has little expected error. Then I want to weight it heavily. So it has a small sigma_2 and that gives it a large weight. And sigma_1 similarly. So that's the weighting for the case that you can actually hope to use in practice.

I'll just mention that statisticians would also say, wait a minute. Measurement two and measurement three might be interconnected. They might not be independent. There might be a covariance. And then that gets them into more great linear algebra actually. But if I want a diagonal matrix C that's the case when my measurements are independent. And basically, I'm whitening the system. I'm making the system white, making it all equal variances by rescaling. By weighting the equations.

Okay, thanks. Wednesday is the next big example of the framework with b and f. See you then.