Okay. This lecture is mostly about the idea of similar matrixes. I'm going to tell you what that word similar means and in what way two matrixes are called similar. But before I do that, I have a little more to say about positive definite matrixes. You can tell this is a subject I think is really important and I told you what positive definite meant -- it means that this -- this expression, this quadratic form, x transpose I x is always positive.

But the direct way to test it was with eigenvalues or pivots or determinants. So I -- we know what it means, we know how to test it, but I didn't really say where positive definite matrixes come from. And so one thing I want to say is that they come from least squares in -- and all sorts of physical problems start with a rectangular matrix -- well, you remember in least squares the crucial combination was A transpose A.

So I want to show that that's a positive definite matrix.

Can -- so I -- I'm going to speak a little more about positive definite matrixes, just recapping -- so let me ask a question.

It may be on the homework. Suppose a matrix A is positive definite. I mean by that it's all -- I'm assuming it's symmetric. That's always built into the definition. So we have a symmetric positive definite matrix. What about its inverse? Is the inverse of a symmetric positive definite matrix also symmetric positive definite? So you quickly think, okay, what do I know about the pivots of the inverse matrix? Not much. What do I know about the eigenvalues of the inverse matrix? Everything, right? The eigenvalues of the inverse are one over the eigenvalues of the matrix.

So if my matrix starts out positive definite, then right away I know that its inverse is positive definite, because those positive eigenvalues -- then one over the eigenvalue is also positive.

What if I know that A -- a matrix A and a matrix B are both positive definite? But let me ask you this. Suppose if A and B are positive definite, what about -- what about A plus B? In some way, you hope that that would be true.

It's -- positive definite for a matrix is kind of like positive for a real number. But we don't know the eigenvalues of A plus B. We don't know the pivots of A plus B. So we just, like, have to go down this list of, all right, which approach to positive definite can we get a handle on? And this is a good one. This is a good one.

Can we -- how would we decide that -- if A was like this and if B was like this, then we would look at x transpose A plus B x. I'm sure this is in the homework. Now -- so we have x transpose A x bigger than zero, x transpose B x positive for all -- for all x, so now I ask you about this guy.

And of course, you just add that and that and we get what we want. If A and B are positive definites, so is A plus B. So that's what I've shown.

So is A plus B. Just -- be sort of ready for all the approaches through eigenvalues and through this expression. And now, finally, one more thought about positive definite is this combination that came up in least squares. Can I do that? So now -- now suppose A is rectangular, m by n.

I -- so I'm sorry that I've used the same letter A for the positive definite matrixes in the eigenvalue chapter that I used way back in earlier chapters when the matrix was rectangular.

Now, that matrix -- a rectangular matrix, no way its positive definite. It's not symmetric.

It's not even square in general.

But you remember that the key for these rectangular ones was A transpose A. That's square.

That's symmetric. Those are things we knew -- we knew back when we met this thing in the least square stuff, in the projection stuff. But now we know something more -- we can ask a more important question, a deeper question -- is it positive definite? And we sort of hope so.

Like, we -- we might -- in analogy with numbers, this is like -- sort of like the square of a number, and that's positive. So now I want to ask the matrix question. Is A transpose A positive definite? Okay, now it's -- so again, it's a rectangular A that I'm starting with, but it's the combination A transpose A that's the square, symmetric and hopefully positive definite matrix.

So how -- how do I see that it is positive definite, or at least positive semi-definite? You'll see that. Well, I don't know the eigenvalues of this product. I don't want to work with the pivots. The right thing -- the right quantity to look at is this, x transpose Ax -- A -- x transpose times my matrix times x.

I'd like to see that this thing -- that that expression is always positive. I'm not doing it with numbers, I'm doing it with symbols. Do you see -- how do I see that that expression comes out positive? I'm taking a rectangular matrix A and an A transpose -- that gives me something square symmetric, but now I want to see that if I multiply -- that if I do this -- I form this quadratic expression that I get this positive thing that goes upwards when I graph it.

How do I see that that's positive, or absolutely it isn't negative anyway? We'll have to, like, spend a minute on the question could it be zero, but it can't be negative.

Why can this never be negative? The argument is -- like the one key idea in so many steps in linear algebra -- put those parentheses in a good way.

Put the parentheses around Ax and what's the first part? What's this x transpose A transpose? That is Ax transpose. So what do we have? We have the length squared of Ax.

We have -- that's the column vector Ax that's the row vector Ax, its length squared, certainly greater than or possibly equal to zero. So we have to deal with this little possibility. Could it be equal? Well, when could the length squared be zero? Only if the vector is zero, right? That's the only vector that has length squared zero.

So we have -- we would like to -- I would like to get that possibility out of there. So I want to have Ax never -- never be zero, except of course for the zero vector. How do I assure that Ax is never zero? The -- in other words, how do I show that there's no null space of A? The rank should be -- so now remember -- what's the rank when there's no null space? By no null space, you know what I mean. Only the zero vector in the null space. So if I have a -- if I have an 11 by 5 matrix -- so it's got 11 rows, 5 columns, when is there no null space? So the columns should be independent -- what's the rank? n 5 -- rank n. Independent columns, when -- so if I -- then I conclude yes, positive definite. And this was the assumption -- then A transpose A is invertible -- the least squares equations all work fine. And more than that -- the matrix is even positive definite.

And I just to say one comment about numerical things, with a positive definite matrix, you never have to do row exchanges.

You never run into unsuitably small numbers or zeroes in the pivot position. They're the right -- they're the great matrixes to compute with, and they're the great matrixes to study. So that's -- I wanted to take this first ten minutes of grab the first ten minutes away from similar matrixes and continue a -- this much more with positive definite.

I'm really at this point, now, coming close to the end of the heart of linear algebra. The positive definiteness brought everything together. Similar matrixes, which is coming the rest of this hour is a key topic, and please come on Monday. Monday is about what's called the SVD, singular values. It's the -- has become a central fact in -- a central part of linear algebra.

I mean, you can come after Monday also, but -- Monday is, -- that singular value thing has made it into this course. Ten years ago, five years ago it wasn't in the course, now it has to be.

Okay. So can I begin today's lecture proper with this idea of similar matrixes.

This is what similar matrixes mean.

So here -- let's start again. I'll write it again.

So A and B are similar. A and B are -- now I'm -- these matrixes -- I'm no longer talking about symmetric matrixes, in -- at least no longer expecting symmetric matrixes. I'm talking about two square matrixes n by n. A and B, they're n by n matrixes. And I'm introducing this word similar. So I'm going to say what does it mean? It means that they're connected in the way -- well, in the way I've written here, so let me rewrite it. That means that for some matrix M, which has to be invertible, because you'll see that -- this one matrix is -- take the other matrix, multiply on the right by M and on the left by M inverse. So the question is, why that combination? But part of the answer you know already. You remember -- we've done this -- we've taken a matrix A -- so let's do an example of similar.

Suppose A -- the matrix A -- suppose it has a full set of eigenvectors. They go in this eigenvector matrix S. Then what was the main point of the whole -- the main calculation of the whole chapter was -- is -- use that eigenvector matrix S and its inverse comes over there to produce the nicest possible matrix lambda. Nicest possible because it's diagonal. So in our new language, this is saying A is similar to lambda.

A is similar to lambda, because there is a matrix, and this particular -- there is an M and this particular M is this important guy, this eigenvector matrix. But if I take a different matrix M and I look at M inverse A M, the result won't come out diagonal, but it will come out a matrix B that's similar to A. Do you see that I'm -- what I'm doing is, like -- I'm putting these matrixes into families. All the matrixes in one -- in the family are similar to each other.

They're all -- each one in this family is connected to each other one by some matrix M and the -- like the outstanding member of the family is the diagonal guy. I mean, that's the simplest, neatest matrix in this family of all the matrixes that are similar to A, the best one is lambda. But there are lots of others, because I can take different -- instead of S, I can take any old matrix M, any old invertible matrix and -- and do it. I'd better do an example.

Okay. Suppose I take A as the matrix two one one two. Okay.

Do you know the eigenvalue matrix for that? The eigenvalues of that matrix are -- well, three and one.

So that -- and the eigenvectors would be easy to find.

So this matrix is similar to this one.

But my point is -- but also, I can also take my matrix, two one one two, I could multiply it by -- let's see, what -- I'm just going to cook up a matrix M here.

I'm -- I'll -- let me just invent -- one four one zero.

And over here I'll put M inverse, and because I happened to make that triangular, I know that its inverse is that, right? So there's M inverse A M, that's going to produce some matrix -- oh, well, I've got to do the multiplication, so hang on a second, let -- I'll just copy that one minus four zero one and multiply these guys so I'm getting two nine one and six, I think.

Can you check it as I go, because you -- see I'm just -- so that's two minus four, I'm getting a minus two nine minus 24 is a minus 15, my God, how did I get this? And that's probably one and six.

So there's my matrix B. And there's my matrix lambda, there's my matrix A and my point is these are all similar matrixes. They all have something in common, besides being just two by two.

They have something in common. And that's -- and what is it? What's the point about two matrixes that are built out of -- the B is built out of M inverse A M. What is it that A and B have in common? That's the main -- now I'm telling you the main fact about similar matrixes. They have the same eigenvalues.

This is -- this chapter is about eigenvalues, and that's why we're interested in this family of matrixes that have the same eigenvalues. What are the eigenvalues in this example? Lambda.

The eigenvalues of that I could compute.

The eigenvalues of that I can compute really fast.

So the eigenvalues are three and one -- for this for sure.

Now did we -- do you see why the eigenvalues are three and one for that one? If I tell you the eigenvalues are three and one, you prick -- quickly process the trace, which is -- and four -- agrees with four and you process the determinant, three times one -- the determinant is three and you say yes, it's right.

Now I'm hoping that the eigenvalues of this thing are three and one. May I process the trace and the determinant for that one? What's the trace here? The trace of this matrix is four minus two and six, and that's what it should be. What's the determinant minus twelve plus fifteen is three. The determinant is three. The eigenvalues of that matrix are also three and one. And you see I created this matrix just like -- I just took any M, like, one that popped into my head and computed M inverse A M, got that matrix, it didn't look anything special but it's -- like A itself, it has those eigenvalues three and one.

So that's the main fact and let me write it down.

Similar matrixes have the same eigenvalues.

So I'll just put that as an important point.

And think why.

Why is that? So that's what that family of matrixes is. The matrixes that are similar to this A here are all the matrixes with eigenvalues three and one. Every matrix with eigenvalues three and one, there's some M that connects this guy to the one you think of.

And then of course, the most special guy in the whole family is the diagonal one with eigenvalues three and one sitting there on the diagonal. But also, I could find -- I mean, tell me just a couple more members of the family.

Another -- tell me another matrix that has eigenvalues three and one. Well, let's see, I -- oh, I'll just make it triangular.

That's in the family. There is some M that -- that connects to this one. And -- and also this.

There's some matrix M -- so that M inverse A M comes out to be that. There's a whole family here. And they all share the same eigenvalues.

So why is that? Okay.

I'm going to start -- the only possibility is to start with Ax equal lambda x. Okay, so suppose A has the eigenvalue lambda. Now I want to get B into the picture here somehow. You remember B is M inverse A M. Let's just remember that over here. B is M inverse A M.

And I want to see its eigenvalues.

How I going to get M inverse A M into this equation? Let me just sort of do it. I'll put an M times an M inverse in there, right? That was -- I haven't changed the left-hand side, so I better not change the right-hand side. So everybody's okay so far, I just put in there -- see, I want to get a -- so now I'll multiply on the left by M inverse -- I have to do the same to this side and that number lambda's just a number, so it factors out in the front. So what I have here is this was safe. I did the same thing to both sides. And now I've got B. There's B. That's B times this vector M inverse x is equal to lambda times this vector M inverse x.

So what have I learned? I've learned that B times some vector is lambda times that vector.

I've learned that lambda is an eigenvalue of B also.

So this is -- if -- so this is -- if lambda's an eigenvalue of A, then I can write it this way and I discover that lambda's an eigenvalue of B. That's the end of the proof.

The eigenvector didn't stay the same.

Of course I don't expect the eigenvectors to stay the same. If all the eigenvalues are the same and all the eigenvectors are the same, then probably the matrix is the same. Here the eigenvector changes, so the eigenvector -- so the point is then the eigenvector of B -- of B is M inverse times the eigenvector of A. Okay.

That's all that this says here. The eigenvector of A was X, and so the M inverse -- similar matrixes, then have the same eigenvalues and their eigenvectors are just moved around. Of course, that's what we -- that's what happened way back -- and the most important similar matrixes are to diagonalize.

So what was the point when we diagonalized? The eigenvalues stayed the same, of course.

Three and one. What about the eigenvectors? The eigenvectors were whatever they were for the matrix A, but then what were the eigenvectors for the diagonal matrix? They're just -- what are the eigenvectors of a diagonal matrix? They're just one zero and zero one.

So this step made the eigenvectors nice, didn't change the eigenvalues, and every time we don't change the eigenvalues.

Same eigenvalues. Okay.

Now -- so I've got all these matrixes in -- I've got this family of matrixes with eigenvalues three and one.

Fine. That's a nice family.

It's nice because those two eigenvalues are different.

I now have to -- to get into that -- the -- into the less happy possibility that the two eigenvalues could be the same. And then it's a little trickier, because you remember when two eigenvalues are the same, what's the bad possibility? That there might not be enough -- a full set of eigenvectors and we might not be able to diagonalize. So I need to discuss the bad case. So the bad -- can I just say bad? If lambda one equals lambda two, then the matrix might not be diagonalizable.

Suppose lambda one equals lambda two equals four, say. Now if I look at the family of matrixes with eigenvalues four and four, well, one possibility occurs to me. One family with eigenvalues four and four has this matrix in it, four times the identity.

Then another -- but now I want to ask also about the matrix four four one zero. And my point -- here's the whole point of this -- of this bad stuff, is that this guy is not in the same family with that one.

The family of a -- of matrixes that have eigenvalues four and four is two families. There's this total loner here who's in a family off -- right? Just by himself.

And all the others are in with this guy.

So the big family includes this one.

And it includes a whole lot of other matrixes, all -- in fact, in this two by two case, it -- you see where -- what do I mean -- so what I using, this word family -- in a family, I mean they're similar. So my point is that the only matrix that's similar to this is itself.

The only matrix that's similar to four times the identity is four times the identity. It's off by itself.

Why is that? The -- if this is my matrix, four times the identity, and I take it, I multiply on the right by any matrix M, I multiply on the left by M inverse, what do I get? This is any M, but what's the result? Well, factoring out a four, that's just the identity matrix in there. So then the M inverse cancels the M, so I've just got this matrix back again. So whatever the M is, I'm not getting any more members of the family.

So this is one small family, because it only has one person.

One matrix, excuse me. I think of these matrixes as people by this point, in eighteen oh six.

Okay, the other family includes all the rest -- all other matrixes that have eigenvalues four and four.

This is somehow the best one in that family.

See, I can't make it diagonal. If I -- if it's diagonal, it's this one. It's in its own, by itself. So I have to think, okay, what's the nearest I can get to diagonal? But it will not be diagonalizable. That -- do you know that that matrix is not diagonalizable? Of

course, because if it was diagonalizable, it would be similar to that, which it isn't. The eigenvalues of this are four and four, but what's the catch with that matrix? It's only got one eigenvector.

That's a non-diagonalizable matrix.

Only one eigenvector. And somehow, if I made that one into a ten or to a million, I could find an M, it's in the family, it's similar. But the best -- so the best guy in this family is this one.

And this is called the Jordan -- so this guy Jordan picked out -- so he, like, studied, these families of matrixes, and each family, he picked out the nicest, most diagonal one. But not completely diagonal, because there's nobody -- there isn't a diagonal matrix in this family, so there's a one up there in the Jordan form.

Okay. I think we've got to see some more matrixes in that family. So, all right, let me -- let's just think of some other matrixes whose eigenvalues are four and four but they're not four times the identity. So -- and I believe that -- that this -- that all the examples we pick up will be similar to each other and -- do you see why -- in this topic of similar matrixes, the climax is the Jordan form.

So it says that every matrix -- I'll write down what the Jordan form -- what Jordan discovered.

He found the best looking matrix in each family.

And that's -- then we've got -- then we've covered all matrixes including the non-diagonalizable one.

That -- that's the point, that in some way, Jordan completed the diagonalization by coming as near as he could, which is his Jordan form.

And therefore, if you want to cover all matrixes, you've got to get him in the picture.

It used to be -- when I took eighteen oh six, that was the climax of the course, this Jordan form stuff.

I think it's not the climax of linear algebra anymore, because -- it's not easy to find this Jordan form for a general matrix, because it depends on these eigenvalues being exactly the same.

You'd have to know exactly the eigenvalues and it -- and you'd have to know exactly the rank and the slightest change in numbers will change those eigenvalues, change the rank and therefore the whole thing is numerically not an -- a good thing. But for algebra, it's the right thing to understand this family. So just tell me another matrix -- a few more matrixes -- so more members of the family. Let me put down again what the best one is. Okay.

All right. Some more matrixes.

Let's see, what I looking for? I'm looking for matrixes whose trace is what? So if I'm looking for more matrixes in the family, they'll all have the same eigenvalues, four and four. So their trace will be eight.

So why don't I just take, like, five and three -- I've got the trace right, now the determinant should be what? Sixteen. So I just fix this up -- shall I put maybe a one and a minus one there? Okay.

There's a matrix with eigenvalues four and four, because the trace is eight and the determinant is sixteen.

And I don't think it's diagonalizable.

Do you know why it's not diagonalizable? Because if it was diagonalizable, the diagonal form would have to be this.

But I can't get to that form, because whatever I do with any M inverse and M I stay with that form.

I could never get -- connect those.

So I can put down more members -- here -- here's another easy one. I could put the four and the four and a seventeen down there. All these matrixes are similar. If I'm -- I could find an M that would show that that one is similar to that one. And in -- you can see the general picture is I can take any a and any 8-a here and any -- oh, I don't know, whatever you put it'd be -- anyway, you can see. I can fill this in, fill this in to make the trace equal eight, the determinant equal 16, I get all that family of matrixes and they're all similar. So we see what eigenvalues do.

They're all similar and they all have only one eigenvector.

So I -- if I'm -- if you were going to -- allow me to add to this picture, they have the same lambdas and they also have the same number of independent eigenvectors.

Because if I get an eigenvector for x I get one for -- for A, I get one for B also. So -- and same number of eigenvectors. But even more than that -- even more than that -- I mean, it's not enough just to count eigenvectors. Yes, let me give you an example why it's not even enough to count eigenvectors.

So another example. So here are some matrixes -- oh, let me make them four by four -- okay, here -- here's a matrix. I mean, like if you want nightmares, think about matrixes like these.

Uh, so a one off the diagonal -- say a one there, how many -- what are the eigenvalues of that matrix? Oh, I mean -- okay. What are the eigenvalues of that matrix? Please. Four 0s, right? So we're really getting bad matrixes now.

So I mean, this is, like -- Jordan was a good guy, but he had to think about matrixes that all -- that had, like -- an eigenvalue repeated four times.

How many eigenvectors does that matrix have? Well, I'm -- eigenvectors will be -- since the eigenvalue is zero, eigenvectors will be in the null space, right? I'm -- eigenvectors have got to be A x equal zero x.

So what's the dimension of the null space? Two. Somebody said two.

And that's right. How -- why? Because you ask what's the rank of that matrix, the rank is obviously two. The number of independent rows is two, the number of independent columns is two, the rank is two so the null -- the dimension of the null space is four minus two, so it's got two eigenvectors. Two eigenvectors. Two independent eigenvectors. All right.

The dimension of the null space is two.

Now, suppose I change this zero to a seven.

The eigenvalues are all still zero, how -- what about -- how many eigenvectors? What's the dimension of the -- what's the rank of this matrix now? Still two, right? So it's okay.

And actually, this would be similar to the one that had a zero in there.

But it's not as beautiful, Jordan picked this one.

He picked -- he put ones -- we have a one on the -- above the diagonal for every missing eigenvector, and here we're missing two because we've got two, so we've got two eigenvectors and two are missing, because it's a four by four matrix.

Okay, now -- but I was going to give you this second example.

0 1 0 0, let me just move the one. Oop, not there. Off the diagonal and zero zero zero zero zero.

Okay. So now tell me about this matrix. Its eigenvalues are four zeroes again. Its rank is two again.

So it has two eigenvectors and two missing.

But the darn thing is not similar to that one.

A -- a count of eigenvectors looks like these could be similar, but they're not.

Jordan -- see, this is like -- a little three by three block and a little one by one block.

And this one is like a two by two block and a two by two block, and those blocks are called Jordan blocks.

So let me say what is a Jordan block? J block number I has -- so a Jordan block has a repeated eigenvalue, lambda I, lambda I on the diagonal.

Zeroes below and ones above. So there's a block with this guy repeated, but it only has one eigenvector. So a Jordan block has one eigenvector only. This one has one eigenvector, this block has one eigenvector and we get two.

This block has one eigenvector and that block has one eigenvector and we get two. So -- but the blocks are different sizes. And that -- it turns out Jordan worked out -- then this is not similar, not similar to this one. So the -- so I'm, like, giving you the whole story -- well, not the whole story, but the main themes of the story -- is here's Jordan's theorem.

Every square matrix A is similar to A Jordan matrix J.

And what's a Jordan matrix J? It's a matrix with these blocks, block -- Jordan block number one, Jordan block number two and so on.

And let's say Jordan block number d.

And those Jordan blocks look like that, so the eigenvalues are sitting on the diagonal, but we've got some of these ones above the diagonal. We've got the number of -- so the number of blocks -- the number of blocks is the number of eigenvectors, because we get one eigenvector per block. So what I'm -- so if I summarize Jordan's idea -- start with any A.

If its eigenvalues are distinct, then what's it similar to? This is the good case.

if I start with a matrix A and it has different eigenvalues -- it's n eigenvalues, none of them are repeated, then that's a diagonal -- diagonalizable matrix -- the Jordan blocks is -- has -- the Jordan matrix is diagonal.

It's lambda. So the good case -- the good case, J is lambda. All -- there are -- d=n. There are n eigenvectors, n blocks, diagonal, everything great.

But Jordan covered all cases by including these cases of repeated eigenvalues and missing eigenvectors.

Okay. That's a description of Jordan.

That -- that's -- I haven't told you how to compute this thing, and it isn't easy.

Whereas the good case is the -- the good case is what 18.06 is about. The -- this case is what 18.06 was about 20 years ago. So you can see you probably won't have on the final exam the computation of a Jordan matrix for some horrible thing with four repeated eigenvalues.

I'm not that crazy about the Jordan form.

But I'm very positive about positive definite matrixes and about the idea that's coming Monday, the singular value decomposition. So I'll see you on Monday, and have a great weekend. Bye.

MIT OpenCourseWare
http://ocw.mit.edu

18.06SC Linear Algebra
Fall 2011

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.