

MIT OpenCourseWare
<http://ocw.mit.edu>

18.06 Linear Algebra, Spring 2005

Please use the following citation format:

Gilbert Strang, *18.06 Linear Algebra, Spring 2005*. (Massachusetts Institute of Technology: MIT OpenCourseWare). <http://ocw.mit.edu> (accessed MM DD, YYYY). License: Creative Commons Attribution-Noncommercial-Share Alike.

Note: Please use the actual date you accessed this material in your citation.

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

MIT OpenCourseWare
<http://ocw.mit.edu>

18.06 Linear Algebra, Spring 2005
Transcript – Lecture 25

-- one and -- the lecture on symmetric matrixes.

So that's the most important class of matrixes, symmetric matrixes. A equals A transpose.

So the first points, the main points of the lecture I'll tell you right away. What's special about the eigenvalues? What's special about the eigenvectors? This is -- the way we now look at a matrix. We want to know about its eigenvalues and eigenvectors and if we have a special type of matrix, that should tell us something about eigenvalues and eigenvectors. Like Markov matrixes, they have an eigenvalue equal one.

Now symmetric matrixes, can I just tell you right off what the main facts -- the two main facts are? The eigenvalues of a symmetric matrix, real -- this is a real symmetric matrix, we -- talking mostly about real matrixes. The eigenvalues are also real.

So our examples of rotation matrixes, where -- where we got E - eigenvalues that were complex, that won't happen now.

For symmetric matrixes, the eigenvalues are real and the eigenvectors are also very special. The eigenvectors are perpendicular, orthogonal, so which do you prefer? I'll say perpendicular. Perp- well, they're both long words. Okay, right.

So -- I have a -- you should say "why?" and I'll at least answer why for case one, maybe case two, the checking the Eigen -- that the eigenvectors are perpendicular, I'll leave to, the -- to the book. But let's just realize what -- well, first I have to say, it -- it could happen, like for the identity matrix -- there's a symmetric matrix.

Its eigenvalues are certainly all real, they're all one for the identity matrix. What about the eigenvectors? Well, for the identity, every vector is an eigenvector.

So how can I say they're perpendicular? What I really mean is the -- they -- this word are should really be written can be chosen perpendicular.

That is, if we have -- it's the usual case.

If the eigenvalues are all different, then each eigenvalue has one line of eigenvectors and those lines are perpendicular here. But if an eigenvalue's repeated, then there's a whole plane of eigenvectors and all I'm saying is that in that plain, we can choose perpendicular ones. So that's why it's a can be chosen part, is -- this is in the case of a repeated eigenvalue where there's some real, substantial freedom.

But the typical case is different eigenvalues, all real, one dimensional eigenvector space, Eigen spaces, and all perpendicular.

So, just -- let's just see the conclusion.

If we accept those as correct, what happens -- and I also mean that there's a full set of them. I -- so that's part of this picture here, that there -- there's a complete set of eigenvectors, perpendicular ones.

So, having a complete set of eigenvectors means -- so normal -- so the usual -- maybe I put the -- usually -- usual -- usual case is that the matrix A we can write in terms of its eigenvalue matrix and its eigenvector matrix this way, right? We can do that in the usual case, but now what's special when the matrix is symmetric? So this is the usual case, and now let me go to the symmetric case. So in the symmetric case, A , this -- this should become somehow a little special.

Well, the λ s on the diagonal are still on the diagonal. They're -- they're real, but that's where they are. What about the eigenvector matrix? So what can I do now special about the eigenvector matrix when -- when the A itself is symmetric, that says something good about the eigenvector matrix, so what is this -- what does this lead to? This -- these perpendicular eigenvectors, I can not only -- I can not only guarantee they're perpendicular, I could also make them unit vectors, no problem, just scale their length to one.

So what do I have? I have orthonormal eigenvectors. And what does that tell me about the eigenvector matrix? What -- what letter should I now use in place of S -- remember S has the eigenvectors in its columns, but now those columns are orthonormal, so the right letter to use is Q . So that's where -- so we've got the letter all set up, so this should be $Q \lambda Q^{-1}$. Q standing in our minds always for this matrix -- in this case it's square, it's -- so these are the columns of Q , of course. And one more thing.

What's Q^{-1} ? For a matrix that has these orthonormal columns, we know that the inverse is the same as the transpose. So here is the beautiful -- there is the -- the great description, the factorization of a symmetric matrix.

And this is, like, one of the famous theorems of linear algebra, that if I have a symmetric matrix, it can be factored in this form.

An orthogonal matrix times diagonal times the transpose of that orthogonal matrix. And, of course, everybody immediately says yes, and if this is possible, then that's clearly symmetric, right? That -- take -- we've looked at products of three guys like that and taken their transpose and we got it back again. So do you -- do you see the beauty of this -- of this factorization, then? It -- it completely displays the eigenvalues and eigenvectors the symmetry of the -- of the whole thing, because -- that product, Q times λ times Q^T , if I transpose it, it -- this comes in this position and we get that matrix back again. So that's -- in mathematics, that's called the spectral theorem.

Spectrum is the set of eigenvalues of a matrix.

Spec- it somehow comes from the idea of the spectrum of light as a combination of pure things -- where our matrix is broken down into pure eigenvalues and eigenvectors -- in mechanics it's often called the principle axis theorem. It's very useful.

It means that if you have -- we'll see it geometrically.

It means that if I have some material -- if I look at the right axis, it becomes diagonal, it becomes -- the -- the directions don't couple together.

Okay. So that's -- that -- that's what to remember from -- from this lecture. Now, I would like to say why are the eigenvalues real? Can I do that? So -- so -- because that -- something useful comes out. So I'll just come back -- come to that question why real eigenvalues? Okay. So I have to start from the only thing we know, $Ax = \lambda x$. Okay. But as far as I know at this moment, λ could be complex. I'm going to prove it's not -- and x could be complex. In fact, for the moment, even A could be -- we could even think, well, what happens if A is complex? Well, one thing we can always do -- this is -- this is like always -- always okay -- I can -- if I have an equation, I can take the complex conjugate of everything.

That's -- no -- no -- so A conjugate x conjugate equal λ conjugate x conjugate, it just means that everywhere over here that there was a -- an i , then here I changed it to $-i$.

That's -- that -- you know that that step -- that conjugate business, that $a+ib$, if I conjugate it it's $a-ib$.

That's the meaning of conjugate -- and products behave right, I can conjugate every factor. So I haven't done anything yet except to say what would be true if, x -- in any case, even if x and λ were complex. Of course, our -- we're speaking about real matrixes A , so I can take that out. Actually, this already tells me something about real matrixes. I haven't used any assumption of A -- A transpose yet. Symmetry is waiting in the wings to be used. This tells me that if a real matrix has an eigenvalue λ and an eigenvector x , it also has -- another of its eigenvalues is $\bar{\lambda}$ with eigenvector \bar{x} . Real matrixes, the eigenvalues come in λ , $\bar{\lambda}$ -- the complex eigenvalues come in λ and $\bar{\lambda}$ pairs.

But, of course, I'm aiming to show that they're not complex at all, here, by getting symmetry in. So how I going to use symmetry? I'm going to transpose this equation to $\bar{x}^T A^T = \bar{x}^T \bar{\lambda}$.

That's just a number, so I don't mind what I put that number. This is -- this is -- this is a -- then again okay.

But now I'm ready to use symmetry.

I'm ready -- so this was all just mechanics.

Now -- now comes the moment to say, okay, if the matrix is symmetric, then this A^T is the same as A .

You see, at that moment I used the assumption. Now let me finish the discussion. Here -- here's the way I finish. I look at this original equation and I take the inner product.

I multiply both sides by -- oh, maybe I'll do it with this one.

I take -- I multiply both sides by x^T .

$x^T A x = \lambda x^T x$. Okay, fine.

All right, now what's the other one? Oh, for the other one I'll probably use this guy.

A- I happy about this? No.

For some reason I'm not. I'm -- I want to -- if I take the inner product of this from the right with x^T , I get $x^T A x = x^T \lambda x$. I- I've done something dumb, because I've got the -- I haven't learned anything.

I've got -- those two equations are identical,¹ so forgive me for doing such a thing, but, I'll look at the book. Okay. So I took the dot product -- ye, somehow, it didn't -- I should've taken the dot product of this guy here with -- that's what I was going to do. $Ax = \lambda x$, right? Okay.

So that -- that was -- that's fine.

That comes directly from that, multiplying both sides by x^T , but now I'd like to get -- why do I have x bars over there? Oh, yes.

Forget this. Okay.

On this one -- right. On this one, I took it like that, I multiply on the right by x^T . That's the idea. Okay.

Now why I happier with this situation now? A proof is coming here. Because I compare this guy with this one. And they have the same left hand side. So they have the same right hand side. So comparing those two, can -- I'll raise the board to do this comparison -- this thing, $x^T A x = \lambda x^T x$ is equal to $x^T \lambda x$. Okay.

And the conclusion I'm going to reach -- I -- I on the right track here? The conclusion I'm going to reach is $\lambda = \lambda$.

I would have to track down the other possibility that this -- this thing is zero, but let me -- oh -- oh, yes, that's important. It's not zero.

So once I know that this isn't zero, I just cancel it and I learn that $\lambda = \lambda$.

And so what can you -- do you -- have you got the reasoning altogether? What does this tell us? λ 's an eigenvalue of this symmetric matrix.

We've just proved that it equaled λ , so we have just proved that λ is real, right? If, if a number is equal to its own complex conjugate, then there's no imaginary part at all. The number is real.

So λ is real. Good.

Good. Now, what -- but it depended on this little expression, on knowing that that wasn't zero, so that I could cancel it out -- so can we just take a second on that one? Because it's an important quantity.

$\bar{x}^T x$. Okay, now remember, as far as we know, x is complex.

So this is -- here -- x is complex, x has these components, x_1, x_2 down to x_n .

And $\bar{x}^T x$, well, it's transposed and it's conjugated, so that's x_1 conjugated x_2 conjugated up to x_n conjugated. I'm -- I'm -- I'm really reminding you of crucial facts about complex numbers that are going to come into the next lecture as well as this one. So what can you tell me about that product -- I -- I guess what I'm trying to say is, if I had a complex vector, this would be the quantity I would -- I would like. This is the quantity I like.

I would take the vector times its transpose -- now what -- what happens usually if I take a vector -- $a^T a$ -- $x^T x$? I mean, that's a quantity we see all the time, $x^T x$.

That's the length of x squared, right? That's this positive length squared, it's Pythagoras, it's x_1^2 plus x_2^2 and so on. Now our vector's complex, and you see the effect? I'm conjugating one of these guys. So now when I do this multiplication, I have $\bar{x}_1 x_1$ and $\bar{x}_2 x_2$ and so on. So this is an -- this is $\sum a+ib$. And this is $\sum a-ib$.

I mean, what's the point here? What's the point -- when I multiply a number by its conjugate, a complex number by its conjugate, what do I get? I get $a^2 - b^2$ -- the imaginary part is gone.

When I multiply $a+ib$ by its conjugate, what's -- what's the result of that -- of each of those separate little multiplications? There's an a^2 and -- and what -- how many -- what's -- b^2 comes in with a plus or a minus? A plus.

i times minus i is a plus b^2 .

And what about the imaginary part? Gone, right? An iab and a minus iab .

So this -- this is the right thing to do.

If you want a decent answer, then multiply numbers by their conjugates. Multiply vectors by the conjugates of x^T . So this quantity is positive, this quantity is positive -- the whole thing is positive except for the zero vector and that allows me to know that this is a positive number, which I safely cancel out and I reach the conclusion. So actually, in this discussion here, I've done two things.

If I reached the conclusion that λ is real, which I wanted to do. But at the same time, we sort of saw what to do if things were complex.

If a vector is complex, then it's $\bar{x}^T x$, this is its length squared. And as I said, the next lecture Monday, we'll -- we'll repeat that this is the right thing and then do the right thing for matrixes and all other -- all other, complex possibilities.

Okay. But the main point, then, is that the eigenvalues of a symmetric matrix, it just -- do you -- do -- where did we use symmetry, by the way? We used it here, right? Let -- can I just -- let -- suppose A was a complex. Suppose A had been a complex number. Could -- could I have made all this work? If A was a complex number -- complex matrix, then here I should have written \bar{A} . I erased the bar because I assumed A was real. But now let's suppose for a moment it's not. Then when I took this step, what should I have? What did I do on that step? I transposed.

So I should have \bar{A} transpose.

In the symmetric case, that was A , and that's what made everything work, right? This -- this led immediately to that.

This one led immediately to this when the matrix was real, so that didn't matter, and it was symmetric, so that didn't matter.

Then I got A . But -- so now I just get to ask you. Suppose the matrix had been complex. What's the right equivalent of sym- symmetry? So the good matrix -- so here, let me say -- good matrixes -- by good I mean real lambdas and perpendicular x -s. And tell me now, which matrixes are good? If they're -- If they're real matrixes, the good ones are symmetric, because then everything went through. The -- so the good -- I'm saying now what's good.

This is -- this is -- these are the good matrixes.

They have real eigenvalues, perpendicular eigenvectors -- good means A equal A transpose if real.

Then -- then that was what -- our proof worked.

But if A is complex, all -- our proof will still work provided \bar{A} transpose is A .

Do you see what I'm saying? I'm saying if we have complex matrixes and we want to say are they -- are they as good as symmetric matrixes, then we should not only transpose the thing, but conjugate it.

Those are good matrixes. And of course, the most important s - the most important case is when they're real, this part doesn't matter and I just have A equal A transpose symmetric. Do you -- I -- I'll just repeat that. The good matrixes, if complex, are these. If real, that doesn't make any difference so I'm just saying symmetric.

And of course, 99% of examples and applications to the matrixes are real and we don't have that and then symmetric is the key property.

Okay. So that -- that's, these main facts and now let me just -- let me just -- so that's this x bar transpose x , okay.

So I'll just, write it once more in this form. So perpendicular orthonormal eigenvectors, real eigenvalues, transposes of orthonormal eigenvectors.

That's the symmetric case, A equal A transpose.

Okay. Good.

Actually, I'll even take one more step here.

Suppose -- I -- I can break this down to show you really what that says about a symmetric matrix.

I can break that down. Let me here -- here go these eigenvectors. I -- here go these eigenvalues, λ_1 , λ_2 and so on.

Here go these eigenvectors transposed.

And what happens if I actually do out that multiplication? Do you see what will happen? There's λ_1 times q_1 transpose. So the first row here is just λ_1 one q_1 transpose. If I multiply column times row -- you remember I could do that? When I multiply matrices, I can multiply columns times rows? So when I do that, I get λ_1 one and then the column and then the row and then λ_2 one and then the column and the row.

Every symmetric matrix breaks up into these pieces. So these pieces have real λ s and they have these Eigen -- these orthonormal eigenvectors.

And, maybe you even could tell me what kind of a matrix have I got there? Suppose I take a unit vector times its transpose? So column times row, I'm getting a matrix. That's a matrix with a special name. What's it's -- what kind of a matrix is it? We've seen those matrices, now, in chapter four. It's -- is $A A^T$ with a unit vector, so I don't have to divide by $A^T A$. That matrix is a projection matrix. That's a projection matrix.

It's symmetric and if I square it there'll be another -- there'll be a q_1 transpose q_1 , which is one. So I'll get that matrix back again. Every -- so every symmetric matrix -- every symmetric matrix is a combination of -- of mutually perpendicular -- so perpendicular projection matrices.

Projection matrices. Okay.

That's another way that people like to think of the spectral theorem, that every symmetric matrix can be broken up that way. That -- I guess at this moment -- first I haven't done an example.

I could create a symmetric matrix, check that it's -- find its eigenvalues, they would come out real, find its eigenvectors, they would come out perpendicular and you would see it in numbers, but maybe I'll leave it here for the moment in letters. Oh, I -- maybe I will do it with numbers, for this reason.

Because there's one more remarkable fact.

Can I just put this further great fact about symmetric matrices on the board? When I have symmetric matrices, I know their eigenvalues are real.

So then I can get interested in the question are they positive or negative? And you remember why that's important. For differential equations, that decides between instability and stability. So I'm -- after I know they're real, then the next question is

are they positive, are they negative? And I hate to have to compute those eigenvalues to answer that question, right? Because computing the eigenvalues of a symmetric matrix of order let's say 50 -- compute its 50 eigenvalues -- is a job.

I mean, by pencil and paper it's a lifetime's job.

I mean, which -- and in fact, a few years ago -- well, say, 20 years ago, or 30, nobody really knew how to do it. I mean, so, like, science was stuck on this problem.

If you have a matrix of order 50 or 100, how do you find its eigenvalues? Numerically, now, I'm just saying, because pencil and paper is -- we're going to run out of time or paper or something before we get it.

Well -- and you might think, okay, get Matlab to compute the determinant of λ minus A , A minus λ I , this polynomial of 50th degree, and then find the roots. Matlab will do it, but it will complain, because it's a very bad way to find the eigenvalues. I'm sorry to be saying this, because it's the way I taught you to do it, right? I taught you to find the eigenvalues by doing that determinant and taking the roots of that polynomial. But now I'm saying, okay, I really meant that for two by twos and three by threes but I didn't mean you to do it on a 50 by 50 and you're not too unhappy, probably, because you didn't want to do it. But -- good, because it would be a very unstable way -- the 50 answers that would come out would be highly unreliable. So, new ways are -- are much better to find those 50 eigenvalues. That's a -- that's a part of numerical linear algebra. But here's the remarkable fact -- that Matlab would quite happily find the 50 pivots, right? Now the pivots are not the same as the eigenvalues. But here's the great thing. If I had a real matrix, I could find those 50 pivots and I could see maybe 28 of them are positive and 22 are negative pivots. And I can compute those safely and quickly. And the great fact is that 28 of the eigenvalues would be positive and 22 would be negative -- that the sines of the pivots -- so this is, like -- I hope you think this -- this is kind of a nice thing, that the sines of the pivots -- for symmetric, I'm always talking about symmetric matrixes -- so I'm really, like, trying to convince you that symmetric matrixes are better than the rest.

So the sines of the pivots are same as the sines of the eigenvalues. The same number.

The number of pivots greater than zero, the number of positive pivots is equal to the number of positive eigenvalues.

So that, actually, is a very useful -- that gives you a good start on a decent way to compute eigenvalues, because you can narrow them down, you can find out how many are positive, how many are negative.

Then you could shift the matrix by seven times the identity.

That would shift all the eigenvalues by seven.

Then you could take the pivots of that matrix and you would know how many eigenvalues of the original were above seven and below seven. So this -- this neat little theorem, that, symmetric matrixes have this connection between the -- nobody's mixing up and thinking the pivots are the eigenvalues -- I mean, the only thing I can think of is the product of the pivots equals the product of the

eigenvalues, why is that? So if I asked you for the reason on that, why is the product of the pivots for a symmetric matrix the same as the product of the eigenvalues? Because they both equal the determinant.

Right. The product of the pivots gives the determinant if no row exchanges, the product of the eigenvalues always gives the determinant.

So -- so the products -- but that doesn't tell you anything about the 50 individual ones, which this does.

Okay. So that's -- those are essential facts about symmetric matrixes.

Okay. Now I -- I said in the -- in the lecture description that I would take the last minutes to start on positive definite matrixes, because we're right there, we're ready to say what's a positive definite matrix? It's symmetric, first of all.

On -- always I will mean symmetric.

So this is the -- this is the next section of the book. It's about this -- if symmetric matrixes are good, which was, like, the point of my lecture so far, then positive, definite matrixes are -- a subclass that are excellent, okay. Just the greatest.

so what are they? They're matrixes -- they're symmetric matrixes, so all their eigenvalues are real. You can guess what they are.

These are symmetric matrixes with all -- the eigenvalues are -- okay, tell me what to write. What -- well, it -- it's hinted, of course, by the name for these things. All the eigenvalues are positive. Okay.

Tell me about the pivots. We can check the eigenvalues or we can check the pivots. All the pivots are what? And then I'll -- then I'll finally give an example.

I feel awful that I have got to this point in the lecture and I haven't given you a single example. So let me give you one.

Five three two two. That's symmetric, fine. It's eigenvalues are real, for sure. But more than that, I know the sines of those eigenvalues.

And also I know the sines of those pivots, so what's the deal with the pivots? The E_i if the eigenvalues are all positive and if this little fact is true that the pivots and eigenvalues have the same sines, then this must be true -- all the pivots are positive.

And that's the good way to test.

This is the good test, because I can -- what are the pivots for that matrix? The pivots for that matrix are five. So pivots are five and what's the second pivot? Have we, like, noticed the formula for the second pivot in a matrix? It doesn't necessarily -- you know, it may come out a fraction for sure, but what is that fraction? Can you tell me? Well, here, the product of the pivots is the determinant. What's the determinant of this matrix? Eleven? So the second pivot must be eleven over five, so that the product is eleven. They're both positive.

Then I know that the eigenvalues of that matrix are both positive. What are the eigenvalues? Well, I've got to take the roots of -- you know, do I put in a minus lambda? You mentally do this -- lambda squared minus how many lambdas? Eight? Right. Five and three, the trace comes in there, plus what number comes here? The determinant, the eleven, so I set that to zero. So the eigenvalues are -- let's see, half of that is four, look at that positive number, plus or minus the square root of sixteen minus eleven, I think five. The eigenvalues -- well, two by two they're not so terrible, but they're not so perfect.

Pivots are really simple. And this is a -- this is the family of matrixes that you really want in differential equations, because you know the sines of the eigenvalues, so you know the stability or not.

Okay. There's one other related fact I can pop in here in -- in the time available for positive definite matrixes. The related fact is to ask you about determinants. So what's the determinant? What can you tell me if I -- remember, positive definite means all eigenvalues are positive, all pivots are positive, so what can you tell me about the determinant? It's positive, too.

But somehow that -- that's not quite enough.

Here -- here's a matrix minus one minus three, what's the determinant of that guy? It's positive, right? Is this a positive, definite matrix? Are the pivots -- what are the pivots? Well, negative.

What are the eigenvalues? Well, they're also the same.

So somehow I don't just want the determinant of the whole matrix. Here is eleven, that's great. Here the determinant of the whole matrix is three, that's positive.

I also -- I've got to check, like, little sub-determinants, say maybe coming down from the left.

So the one by one and the two by two have to be positive. So there -- that's where I get the all. All -- can I call them sub-determinants -- are -- see, I have to -- I need to make the thing plural. I need to test n things, not just the big determinant. All sub-determinants are positive. Then I'm okay.

Then I'm okay. This passes the test.

Five is positive and eleven is positive.

This fails the test because that minus one there is negative. And then the big determinant is positive three. So t- this -- these -- this fact -- you see that actually the course, like, coming together. And that's really my point now.

In the next -- in this lecture and particularly next Wednesday and Friday, the course comes together.

These pivots that we met in the first week, these determinants that we met in the middle of the course, these eigenvalues that we met most recently -- all matrixes are square here, so coming together for square matrixes means these three pieces come together and they come together in that beautiful fact, that if -- that all the -- that if

I have one of these, I have the others. That if I -- but for symmetric matrixes. So that -- this will be the positive definite section and then the real climax of the course is to make everything come together for n by n matrixes, not necessarily symmetric -- bring everything together there and that will be the final thing. Okay.

So have a great weekend and don't forget symmetric matrixes.

Thanks.