GILBERT STRANG:
So I've worked hard over the weekend. I figured out what I was doing last time and what I'm doing this time and improved the notes. So you'll get a new set of notes on the last lecture and on this one. And I kind of got a better picture of what we're doing. And that board is aiming to describe the large picture of what we're doing last time and this time.

So last time was about changes in A inverse when A changed. This time is about changes in eigenvalues and changes in singular values when A change. As you can imagine, this is a natural important situation. Matrices move, and therefore, their inverses change, their eigenvalues change, their singular values change. And you hope for a formula.

Well, so we did have a formula for last time for the change in the inverse matrix. And I didn't get every u and v transpose in the right place in the video or in the first version of the notes, but I hope that that formula, that that Woodbury Morrison formula will be correct this time. So I won't go back over that part.

But I realize also there is another question that we can answer when the change is very small, when the change in A is dA or delta A, a small change. And that's, of course, what calculus is about. So I have to sort of parallel topics here. What is the derivative when the change is infinitesimal? And what is the actual change when the change is finite size?

So now, let me say what we can do and what we can't do. Oh, I'll start out by figuring out what the derivative is for the inverse. So that's like completing the last time for infinitesimal changes. Then I'll move on to changes in the eigenvalues and singular values. And there, you cannot expect an exact formula. We had a formula that was exact, apart from any typos, for this. And we'll find a formula for this, and we'll find a formula for that and for that. Well, that one will come from this one. So this will be a highlight today. How do the eigenvalues change when the matrix changes?

But we won't be able to do parallel to this, we won't be able to-- oh, we will be able to do something for finite changes. That's important. Mathematics would have to keep hitting that

problem until it got somewhere. So I won't get an exact formula for that change. That's too much.

But I'll get inequalities. How big that change could be. What can I say about it? So these are highly interesting.

May I start with completing the last lecture? What is the derivative of the inverse? So I'm thinking here, so what's the setup? The setup is my matrix A depends on time, on t. And it has an inverse. A inverse depends on t.

And if I know this dependence, in other words, if I know dA dt, how the matrix is depending on t, then I hope I could figure out what the derivative of A inverse is. We should be able to do this. So let me just start with-- it's not hard and it complements this one by doing the calculus case, the infinitesimal change. So I want to get to that.

I can figure out the change in A. And my job is to find the derivative of A inverse. So here's a handy identity. Can I just put this here? So here's my usual identity.

So as last time, I start with a finite change because calculus always does that, right. It starts with a delta t and then it goes to 0. So here I am up at a full size change. So I think that this is equal to B inverse A minus B A inverse. And if it's true, it's a pretty cool formula.

And, look, it is true, because over on this right-hand side, I have B inverse times A A inverse. That's the identity. So that's my B inverse. And I have the minus, the B inverse B is the identity. There's A inverse. It's good, right?

So from that, well, I could actually learn from that the rank of this equals the rank of this. That's a point that I made from the big formula. But now, we can see it from an easy formula. Everywhere here, I'm assuming that A and B are invertible matrices. So when I multiply by an invertible matrix, that does not change the rank. So those have the same ranks.

But I want to get further than that. I want to find this. So how do I go? How do I go forward with that job to find the derivative of the inverse?

Well, I'm going to call this a change in A inverse. And over here, I'll have B will be-- yeah, OK, let's see, am I right? Yeah.

So B inverse will be-- this is A plus delta A inverse. And this is-- well, that's A minus B. So that's

really minus delta A. From A to B is the change. Here, I'm looking at the difference A minus B. So it's minus a change. And here, I have A inverse. I haven't done anything except to introduce this delta and get B out of it and brought delta in.

Now, I'm going to do calculus. So I'm thinking of B as there's a sort of a delta t. And I'm going to divide by both sides by delta t. I have to do this if I want-- and now, I'll let delta t go to 0. So calculus appears. Finally, our-- I won't say enemy calculus, but there is a sort of like competition between linear algebra and calculus for college mathematics. Calculus has had far, far too much time and attention. It like it gets three or four semesters of calculus for people who don't get any linear algebra. I'm glad this won't be on the video, but I'm afraid it will.

Anyway, of course, calculus is fine in its place. So here's its place. Now let delta t go to 0. So what does this equation become?

Then everybody knows that as the limit of delta t goes to 0, I replace deltas by-- so this delta A divided by delta t that has a meaning. The top has a meaning and the bottom has for me. But then the limit, it's the ratio that has a meaning.

So dA by itself, I don't attach a meaning to that. That's infinitesimal. It's the limit, so that's why I wanted a delta over a delta so I could do calculus. So what happens now is delta t goes to 0. And, of course, as delta t goes to 0, that carries delta A to 0. So that becomes A inverse.

And what does this approach as delta t goes to 0? dA dt with that minus sign. Oh, I've got to remember the minus sign. The minus sign is in here. So I'm bringing out the minus sign.

Then this was A inverse, as we had. And that's dA dt. And that's A inverse. That's our formula, a nice formula, which sort of belongs in people's knowledge.

You recognize that if A was a 1 by 1 matrix, we could call it x, instead of A. If A was a 1 by 1 matrix x, then I'm saying the formula for the derivative of 1 over x, right? A inverse just 1 by 1 case is just 1 over x. So the derivative of 1-- or maybe t, I should be saying.

If A is just t, then the derivative of 1 over t with respect to t is....? Is minus 1 over t squared. The 1 by 1 case we know. That's what calculus does. And now we're able to do the n by n case. So that's just like good.

And then it's sort of parallel to formulas like this, where this delta A has not gone to 0. It's full size, but low rank. That was the point. Actually, the formula would apply if the rank wasn't low.

But the interest is in low rank here.

Are we good for this? That's really the completion of last time's lecture with derivatives.

OK, come back to here, to the new thing now, lambdas. Let's focus on lambdas, eigenvalues. How does the eigenvalue change when the matrix changes? How does the eigenvalue change when the matrix changes?

So I have two possibilities. One is small change when I'm doing calculus and I'm letting a delta t go to 0. The other is full size, order 1 change, where I will not be able to give you a formula for the new lambdas, but I'll be able to tell you important facts about them. So this is today's lecture now. You could say that's the completion of Friday's lecture.

What about d lambda dt? It's a nice formula. Its proof is fun too. I was very happy about this proof. OK, so I guess calculus is showing up here on this middle board.

So how do I start with the eigenvalues? Well, start with what I know. So these are facts, you could say, that I have to get the eigenvalues into it. And, of course, eigenvalues have to come with eigenvectors.

So I'll again use A of t. It will be depending on t. And an eigenvector that depends on t is an eigenvalue that depends on t times and eigenvector that depends on t. Good? That's fact one that we plan to take the derivative of somehow.

There's also a second fact that comes into play here. What's the deal on the eigenvalues of A transpose? They are the same. The eigenvalues of A transpose are the same as the eigenvalues of A.

Are the eigenvectors the same? Not usually. Of course, if the matrix was symmetric, then A and A transpose are just the same thing. So A transpose would have that eigenvalue-- eigenvector.

But, generally, it has a different eigenvector. And really to keep a sort of separate from this one, let's call it y. It will have the same eigenvalue. I'm going to call it y. But I'm going to make it a row vector, because A transpose is what-- instead of writing down A transpose, I'm going to stay with A, but put the eigenvalue on the left side.

So here's is the eigenvalue-- eigenvector for A on the left. And it has the same eigenvalue

times that eigenvector. But that eigenvector is a row eigenvector, of course. This is an equality between rows. A row times my matrix gives a row. So that's the eigenvalues of-- and it has the same eigenvalues.

So this is totally parallel to that, totally parallel. And maybe sort of less-- definitely less seen, but it's just the same thing for A transpose. Everybody sees that if I transpose this equation, then I've got something that looks like that. But I'd rather have it this way.

Now, one more fact I need. There is-- there has to be some normalization. What should be the length of these? Right now, x could have any length. y could have any length. And there's a natural normalization, which is y transpose times x equal to 1. That normalizes the two. It doesn't tell me the length of x or the length of y. But it tells me, the key thing, the length of both.

So what I've got there is tracking along one eigenvalue and its pair of eigenvectors. And you're always welcome to think of the symmetric case when y and x are the same. And then I would call them q. Oh, well, I would call them q if it was a symmetric matrix. So if it's a symmetric matrix, both eigenvectors would be called q. And this would be saying that q is a--

AUDIENCE: Unit vector.

GILBERT STRANG: Unit vector, right. So this is all stuff we know. And actually, maybe I should write it in matrix notation, because it's important. That's for one eigenvector. This is for all of them at once. Everybody's with it? The x's are the columns of x. And lambda is the diagonal matrix of lambdas. And it has to sit on the right so that it will multiply those columns. So this is like all eigenvectors at once.

What would this one be? This would be like y transpose A equals A--

AUDIENCE: y transpose inverse?

GILBERT STRANG: y transpose, yes-- equals-- and probably these are multiplied-- I feel wrong if I write y transpose here. Like here, the x was on the right and on the left. And I'll-- oh, yeah, y transpose, yeah. OK, so what do I put? Lambda y transpose. Thanks.

And what do I put here? What does this translate to if this was for one eigenvector? For all of them at once, it's just going to translate to y transpose x equal the identity.

This is pretty basic stuff. But stuff somehow we don't always necessarily see. Those are the key facts.

And now, I plan to take the derivative, take the derivative of respect to lambda. Oh, I can derive one more fact. So this would be a formula. This is formula 1. Formula 1 just says, what do I get if I hit this on the left by y transpose? Can I do that? y transpose of t A of t x of t equals lambda of t. That's a number. So I can always bring that out in front of the inner product of vector notation.

Are you good for that? I'm pleading like everything I've done is totally OK. And now, I have a improvement to make on this right-hand side, which is...? So what is y transpose times x?

AUDIENCE: 1.

GILBERT STRANG: It's 1. So let's remember that. It's 1. So in other words, I have got a formula for lambda of t. As time changes, the matrix changes. Its eigenvalues change according to this formula. Its eigenvectors change according of this formula. And its left eigenvectors change according to that formula. So everything here is above board.

And now, what's the point? The point is I'm going to find this, the derivative. So I'm going to take the derivative of that equation and see what I get. That'll be the formula for the derivative of an eigenvalue. And amazingly, it's not that widely known. Of course, it's classical, but it's not always part of courses. So this is as time varies, the matrix varies, A. And therefore, its eigenvalues vary, and its eigenvectors vary.

So we're going to find d lambda dt. It's one level of difficulty more to find dx dt, the derivative of the eigenvector or the second derivative of the eigenvalue. Those kind of come together. And I'm not going to go there. I'm just going to do the one great thing here-- take the derivative of that equation.

Shall I do it over there? So here we go. So I want to compute d lambda dt. And I'm using this formula for lambda there.

So I've got three things that depend on t. And I'm taking the derivative of their product. So I'm going to use the product rule. I'll apply the product rule to that derivative.

Take the derivative of the first guy times A times x. Take the derivative of the second guy times the second guy and the third guy. y transpose of t A of t dx dt. OK? We are one minute

away from a great formula. And I'm really happy if you allow me to say it. That that formula comes by just taking those facts we know, putting them together into this expression that we also know, and this is like lambda equals x inverse Ax. That's a diagonalizing thing and then taking the derivative.

So what do I get if I take that derivative? Well, this term I'm going to keep. I'm not going to play with that. Everybody is clear? That's a number. Here's a matrix. dA dt is a matrix. I take the derivative of every entry in A.

Here's its column vector, its eigenvector. And here's a row vector. So row times matrix times column is a number, 1 by 1. And actually, that's my answer. That's my answer.

So I'm saying that these two terms cancel each other out as those two terms added to zero. This is the right answer for the derivative. That's a nice formula.

So to find the derivative of an eigenvalue, the matrix is changing, you multiply by the eigenvector and by the left eigenvector. It gives you a number. And that's the d lambda dt.

So why do those two guys add to 0? That's all that remains here. And then this topic is ended with this nice formula. So I want to simplify that, simplify that, and show that they cancel each other.

So what is Ax? It's lambda x. So this guy is nothing but it's lambda that depends on time of course times dy dt dy transpose dt. I'm just copying that. Ax is lambda x. Sorry, I didn't mean to make that look hard.

You OK with that? Ax is lambda x. And I am perfectly safe, because lambda is just a number to bring it out to the left. So it doesn't look like it's in the way.

And what about this other term? So I have y transpose-- oh, y transpose A, what's that? What's y transpose A? That's the combination that I know. y transpose A, y is that left eigenvalue. y transpose A brings out a lambda. So this also brings out a lambda times y transpose times dx dt. OK? I just use Ax equal lambda x there. It was really nothing.

Now, what do I do? I want this to be 0. Can you see it happening? It's a great pleasure to see it happening.

So what do I have here? What's my first step now?

AUDIENCE: Like take lambda--

GILBERT STRANG: Bring lambda outside. That's not 0. We don't know what that is. Bring lambda outside there times the whole thing.

So for some wonderful reason I believe that this number, which is a row times a column, a row times a column, two terms there, I believe they knock each other out and that result is 0. And why? Why? Because I come back to-- this board has all that I know. And here's y transpose times x equal 1.

And how does that help me? Because what I'm seeing in that square, in those brackets is?

AUDIENCE: The derivative of y transpose--

GILBERT STRANG: The derivative of the y transpose x. So it's the derivative of?

AUDIENCE: 1

GILBERT STRANG: 1. Therefore, 0. So this is the derivative of 1. It equals 0. Those two terms knock each other out and leave just the nice term that we're seeing.

So the derivative of the eigenvalue, just to have one more look at it before we leave it. The derivative of the eigenvalue is this formula. It's the rate at which the matrix is changing times the eigenvectors on right and left. Sometimes they're called the right eigenvector and the left eigenvector at the time t.

So we're not saying in this d lambda dt. In other words, I get a nice formula, which doesn't involve the derivative of the eigenvector. That's the beauty of it. If I want to go up to take the next step-- I tried this weekend, but it's a mess.

It would be to take the-- so this is my formula then, d lambda dt equals this. And I can take the next derivative of that, and it will involve d second dt squared. But it will also involve dx dt and dy dt. And in fact, a pseudo inverse even shows up. It's another step, and I'm not going that far, because we've got the best formula there.

So now that has answered this question. And I could answer that question the same way. It would involve A transpose A and the singular vectors, instead of involving A and the eigenvectors. Maybe that's a suitable exercise. I don't know. I haven't done it myself.

What I want to do is this, now say, what can we say about the change in the eigenvalue-- and I'll just stay first of all with eigenvalue-- when the change is like rank 1? This is a perfect example when the change is rank 1.

So what can we say about the eigenvalues-- let's take the top, the largest eigenvalue, or all of them, all of them, lambda j, all of them-- of A plus a rank 1 matrix uv transpose. Oh, let's do the nice case here, the nice case, because if I allow a general matrix A, I have to worry about does it have enough eigenvectors? Can it diagonalize? All that stuff. Let's make it a symmetric matrix. And let's make the rank 1 change symmetric too.

So the question is, what can I say about the eigenvalues after a rank 1 change? So again, this isn't calculus now, because the change that I'm making is a true vector and not a differential. And I'm not going to have an exact formula for the new eigenvalues, as I said.

But what I am going to do is write down the beautiful facts that are known about that. And here they are. So, first of all, the eigenvalues are in descending order. We use descending order for singular values. Let's use them also for eigenvalues. So lambda 1 is greater or equal to lambda 2, greater or equal to lambda 3, and so on.

Oh, give me-- give me an idea. What do you expect from that rank 1 change? So that change is rank 1. Can you tell me any more about that change, u u transpose? What kind of a matrix is u u transpose? It's rank 1, but we can say more. It is...?

AUDIENCE: Symmetrical.

GILBERT STRANG: Symmetric, of course. And it is...? Yeah?

AUDIENCE: Positive semidefinite.

GILBERT STRANG: Positive semidefinite. Positive semidefinite. This is a positive change. u u transpose is the typical rank 1 positive semidefinite. It couldn't be positive definite, because it's only got rank 1.

What's the eigenvector of that matrix? Let's just-- why not here? We can do this in two seconds. So u u transpose, that's the matrix I'm asking you to think about. And it's a full n by n matrix, column times a row. Tell me an eigenvector of that matrix. Yes?

AUDIENCE: u.

GILBERT STRANG:	u. If I multiply my matrix by u, I get-- what do I get? I get some number times u. And what is that number lambda?
AUDIENCE:	u transpose u.
GILBERT STRANG:	That lambda happens to be u transpose u. So that's different from u u transpose. This is a matrix. This is 18.065 now. That's a number. And what can you tell me about that number? It is...?
AUDIENCE:	Greater than or equal to 0.
GILBERT STRANG:	Greater-- well, even more. Greater than 0. Greater, because this is a true vector here. So this is greater than 0. It's the only eigenvalue-- all the other eigenvalues of that rank 1 matrix are zero. But the one non-zero eigenvalue is over on the plus side. It's u transpose u. We all recognize that as the length of u squared. It's certainly positive. So we do have a positive semidefinite definite matrix.

What would your guess be of the effect on the eigenvalues of A? So I'm coming back to my real problem-- eigenvalues of S, sorry, S. Symmetric matrices, I'm saying symmetric.

What is your guess if I have a symmetric matrix and I add on u u transpose? What do you imagine that does to the eigenvalues? You're going to get it right. Just say it. What happens to the eigenvalues of S if I add on u u transpose? They will...? |
AUDIENCE:	More positive.
GILBERT STRANG:	They'll be more positive. They'll go up. This is a positive thing. It's like adding 17 to something. It moves up. So therefore, what I believe is-- so I've got two sets of eigenvalues now. One is the eigenvalues of s. The other is the different eigenvalues of S. So I can't call them both lambdas or I'm in trouble. So do you have a favorite other Greek letter for the eigenvalues of S?
AUDIENCE:	Gamma.
GILBERT STRANG:	Gamma. OK, gamma. As long as you say a Greek letter that I have some idea how to write. Zeta, it seems to me, is like the world's toughest letter to write. And electrical engineers can coolly flush off a zeta. I've never succeeded. So I'll write-- what did you say?

AUDIENCE: Gamma.

GILBERT STRANG: Gamma j of the original. So those are the eigenvalues of the original. These are the eigenvalues of the modified. And we're expecting the lambdas to be bigger than the gammas. So that's just a qualitative statement. And it's true. Each lambda is bigger than the gamma. Sorry, yeah, yeah, each lambda, by adding this stuff, the lambdas are bigger than-- so I'll just write that. Lambdas are bigger than gammas.

And that's a fundamental fact, which we could prove. But a little more is known. Of course, the question is, how much bigger? How much can they be way bigger? Well, I don't believe they could be bigger by more than that number myself. But there's just better news than that.

So the lambdas are bigger than the gammas. So lambda 1 is bigger than gamma 1. So this is the S plus u u transpose matrix. And these are the eigenvalues of the S matrix. Lambda 1 is bigger than gamma 1.

But look what's happening in this line of text here. I'm saying that gamma 1-- that lambda 2 is smaller than gamma 1. Isn't that neat? The eigenvalues go up. But they don't just like go anywhere. And that's called interlacing.

So this is one of those wonderful theorems that makes your heart happy, that if I do I rank one change and it's a positive change, then the eigenvalues increase, but they don't increase-- the new eigenvalue is below the new second eigenvalue. It doesn't pass up the old, first eigenvalues. And the new third eigenvalue doesn't pass up the old second eigenvalue. So that's the interlacing theorem that's associated with the names of famous math guys. And of course you have to say that's beautiful.

While we're writing down such a theorem, make a guess of what the theorem would be if I do a rank 2 change. Suppose I do an S, staying symmetric. And I do a rank 1 change. But then I also do a rank 2 change, say w w transpose.

So what's the deal here? What do I know about the change matrix, the delta S here? I know its rank is 2. I'm assuming u and w are not in the same direction. So that's a rank 2 matrix.

And what can you tell me about the eigenvalues of that rank 2 matrix? So it's got n eigenvalues because it's an n by n matrix. But how many non-zero eigenvalues has it got? Two, because its rank is 2. The rank tells you the number of non-zero eigenvalues when matrices are symmetric. It doesn't tell you enough. If matrices are unsymmetric, eigenvalues

can be weird. So stay symmetric here.

So this has two non-zero eigenvalues. And can you tell me their sign. Is that matrix positive semidefinite? Yes, of course, it is. Of course. So this was and this was. And together it certainly is.

So now, I've added a rank 2 positive semidefinite matrix. And now, I'm not going to rewrite this line, but what would you expect to be true? You would expect that the eigenvalues increase. But how big could gamma-- yeah, so gamma 2, let's follow gamma 2.

Well, maybe I should use another-- do the Greeks have any other letters than lambda and gamma. They must had--

AUDIENCE: Zeta.

GILBERT STRANG: Who? C? Hell with that. Who knows one I can write?

AUDIENCE: Alpha.

GILBERT STRANG: Alpha. Good, alpha. Yes, alpha. Right. So alpha is the eigenvalues of this rank 2 change. OK. Now, what am I going to be able to say? Can I say anything about the-- well, of course, alpha 1 is bigger than lambda 1, which was bigger than-- eigenvalues are going up, right? I'm adding positive definite or positive semidefinite stuff. There's no way eigenvalues can start going down on me.

So alpha 1 is a greater or equal to the lambda 1, which had just a rank 1 change, which is greater or equal to the-- mu, was it mu?

AUDIENCE: Gamma.

GILBERT STRANG: Gamma. Gamma 1, and so on. OK, now, let's see, is gamma 1 bigger than alpha-- what am I struggling to write down here? What could I say? Well, what can I say that reflects the fact that this lambda 2-- or sorry, so gamma 1 went up. Gamma 1 was bigger than lambda 2. That was the point here. Gamma 1 is bigger.

So this was a sort of easy, because I'm adding stuff. I expected the lambda to go up. This is where the theorem is that it didn't go up so far as to pass-- or sorry, the lambda 2, which went up, didn't pass up gamma 1. Lambda 2 didn't pass up gamma 1.

And now let me write those words down. Now the alpha 2-- well, could alpha 2 pass up lambda 1? And what about alpha 3?

Let me say what I believe. I think alpha 2, which is like 1 behind, but I'm adding rank 2, I think alpha 2 could pass up lambda 1. It could pass lambda 1. But alpha 3 can't. I believe that alpha 3 is smaller than lambda 1-- smaller than gamma 1, the original. Got it. Yeah, yeah, yeah.

Anyway, I'll get it right in the notes. You know what question I'm asking. And for me, that's the important thing.

Now, there is a little matter of why is this true? This is the good case. Let me give you another example of interlacing. Can I do that? It really comes from this, but let me give you another example that's just striking.

So I have a symmetric matrix, n by n. Call it S. And then I throw away the last row and column. So in here is S n minus 1. The big matrix was Sn. This one is of size n minus 1. So it's got sort of less degrees of freedom, because the last degree of freedom got removed.

And what do you think about the eigenvalues of the n minus 1 eigenvalues of this and the n eigenvalues of that? They interlace. So this has eigenvalue lambda 1. This would have an eigenvalue smaller than that. This would have an eigenvalue lambda 2. This would have an eigenvalue smaller than that and so on.

Just the same interlacing and basically for the same reason, that when you-- this reduction to size n minus 1 is like I'm saying xn has to be 0 in the energy or any of those expression. And the fact of making xn be 0 is like one constraint, taking one degree of freedom away. It reduces the eigenvalues, but not by two.

OK, now I have one final mystery. And let me try to tell you what. It worried me. Now what is it that worried me? Yes, suppose this change, this u, this change that I'm making, suppose it's actually the second eigenvector of S. So can I write this down?

Suppose u is actually the second eigenvector of S. What do I mean by that? So I mean that S times u is lambda 2 times u.

Now, I'm going to change it. S plus u u transpose, that's what I've been looking at. And that moves the eigenvalues up. But what worries me is like if I multiply this by 20, some big

number, I'm going to move that eigenvalue way up, way past. I got worried about this inequality.

When I add this, that same u is lambda 2 plus 20 u. Su is lambda 2u. And this 20 is 20. u is a unit vector.

So you see my worry? Here, I'm doing a rank 1 change. But it's moved an eigenvalue way, way up. So how could this statement be true? So I've just figured out here what gamma-- well, do you see my question? I could leave it as a question to answer next time. Let me do that. And I'll put it online so you'll see it clearly.

It looks like and it happens this eigenvector now has eigenvalue lambda 2 plus 20. Why doesn't that blow away this statement? I'll put that, because it's sort of coming with minus 10 seconds to go in the class, so let's leave that and a discussion of this for next time. But I'm happy with this lecture if you are. Last lecture I got u's and v's mixed up. And it's not reliable. Here, I like the proof of the lambda dt and we're started on this topic 2. Good. Thank you.