

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high-quality educational resources for free. To make a donation or to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

JAMES W. SWAN: Let's go ahead and get started. I hope everybody saw the correction to a typo in homework 1 that was posted on Stellar last night and sent out to you. That's going to happen from time to time. We have four course staff that review all the problems. We try to look through it for any issues or ambiguities. But from time to time, we'll miss something and try to make a correction.

The TAs gave a hint that would have let you solve the problem as written. But that's more difficult than what we had intended for you guys. We don't want to give you a homework assignment that's punishing. We want to give you an assignment that'll help you learn.

Some people in this class that are very good at programming have apparently already completed that problem with the hint. But it's easier, as originally intended. And the correction resets that.

So maybe you'll see the distinction between those things and understand why one version of the problem is much easier than another. But we try to respond as quickly as possible when we notice a typo like that so that we can set you guys on the right course.

So we've got two lectures left discussing linear algebra before we move on to other topics. We're still going to talk about transformations of matrices.

We looked at one type of transformation we could utilize for solving systems of equations. Today, we'll look at another one, the eigenvalue decomposition. And on Monday, we'll look at another one called the singular value decomposition.

Before jumping right in, I want to take a minute and see if there are any questions that I can answer, anything that's been unclear so far that I can try to reemphasize or focus on for you. I was told the office hours are really well-attended. So hopefully, you're getting an opportunity to ask any pressing questions during the office hours or you're meeting with the instructors after class to ask anything that was unclear.

We want to make sure that we're answering those questions in a timely fashion. This course

moves at a pretty quick pace. We don't want anyone to get left behind.

Speaking of getting left behind, we ran out of time a little bit at the end of lecture on Wednesday. That's OK. There were a lot of good questions that came up during class.

And one topic that we didn't get to discuss is formal systems for doing reordering in systems of equations. We saw that reordering is important. In fact, it's essential for solving certain problems via Gaussian elimination.

You won't be able to solve them. Either you'll incur a large numerical error because you didn't do pivoting-- you'd like to do pivoting in order to minimize the numerical error-- or you need to reorder in order to minimize fill-in.

As an example, I've solved a research problem where there was something like 40 million equations and unknowns, a system of partial differential equations. And if you reorder those equations, then you can solve via Gaussian elimination pretty readily. But if you don't, well-- my PC had-- I don't know-- like, 192 gigabytes of RAM.

The elimination on that matrix will fill the memory of that PC up in 20 minutes. And you'll be stuck. It won't proceed after that. So it's the difference between getting a solution and writing a publication about the research problem you're interested in and not.

So how do you do reordering? Well, we use a process called permutation. There's a certain class of matrix called a permutation matrix that can-- its action, multiplying another matrix, can swap rows or columns.

And here's an example of a permutation matrix whose intention is to swap row 1 and 2 of a matrix. So here, it looks like identity, except rather than having 1, 1 on the first two elements of the diagonal, I have 0, 1 and 1, 0.

Here's an example where I take that sort of a matrix, which should swap rows 1 and 2, and I multiply it by a vector. If you do this matrix vector multiplication, you'll see initially, the vector was x_1, x_2, x_3 . But the product will be x_2, x_1, x_3 . It swapped two rows in that vector. Of course, a vector is just a matrix, right? It's an N by 1 matrix.

So P times A is the same as a matrix whose columns are P times each of the columns of A . That's what this notation indicates here. And we know that P times a vector, which is the column from A , will swap two rows in A , right? So the product here will be all the rows of A , the

different rows of A^T with row 1 and 2 swapped with each other.

So permutation, multiplication by the special type of matrix, a permutation matrix, does reordering of rows. If I want to swap columns, I multiply my matrix from the right, AP^T .

So if I want to swap column 1 and 2, I multiply A from the right by P^T . How can I show that that swaps columns? Well, A times P^T is the same as P times A^T . P swaps rows. So it's swapping rows of A^T , which is like swapping columns of A .

So we had some identities associated with matrix-matrix multiplication and their transposes. And you can use that to work out how this permutation matrix will swap columns instead of rows if I multiply from the right instead of the left.

Here's an important concept to know. Permutation matrices are-- would refer to as unitary matrices. They're transposed. It's also they're inverse.

So P times P^T is identity. If I swap the rows and then I swap them back, I get back what I had before. So there are lots of matrices that have this property that they're unitary. We'll see some today.

But permutation matrices are one class, maybe the simplest class, of unitary matrices. They're just doing row or column swaps, right? That's their job.

And so if I have some reordering of the equations or rows of my system of equations that I want, that's going to be indicated by a permutation matrix-- say, P_1 . And I would multiply my entire system of-- both sides of my system of equations by P_1 . That would reorder the rows.

If I have some reordering of the columns or the unknowns in my problem, I would use a similar permutation matrix, P_2 . Of course, P_2^T times P_2 is identity. So this product here does nothing to the system of equations. It just swaps the unknown.

So there's a formal system for doing this sort of swapping. There are a couple other slides that are in your notes from last time that you can look at and I'm happy to answer questions on. We don't have time to go into detail. It discusses the actual methodology, the simplest possible methodology, for doing this kind of reordering or swapping.

So this is a form of preconditioning. If it's preconditioning for pivoting, it's designed to minimize

numerical error. If it's preconditioning in order to minimize fill-in instead, that's meant to make the problem solvable on your computer. But it's a form of preconditioning a system of equations. And we discussed preconditioning before.

So now we know how to solve systems of equations. It's always done via Gaussian elimination if we want an exact solution. There are lots of variants on Gaussian elimination that we can utilize. You're studying one of them in your homework assignment now, where you know the matrix is banded with some bandwidth.

So you don't do elimination on an entire full matrix. You do it on a sparse matrix whose structure you understand. We discussed sparse matrices and a little bit about reordering and now permutation. I feel like my diffusion example last time wasn't especially clear. So let me give you a different example of diffusion.

You guys know Plinko? Have you seen *The Price Is Right*? This is a game where you drop a chip into a board with pegs in it. It's a model of diffusion.

The Plinko chip falls from level to level. It hits a peg. And it can go left or it can go right with equal probability. So the Plinko chip diffuses as it falls down. This guy's excited.

[LAUGHTER]

He just won \$10,000.

[LAUGHTER]

There's a sparse matrix that describes how the probability of finding the Plinko chip in a certain cell evolves from level to level. It works the same way the cellular automata model I showed you last time works.

If the chip is in a particular cell, then at the next level, there's a 50/50 chance that I'll go to the left or I'll go to the right. It looks like this, right?

If the chip is here, there's a 50/50 chance I'll go here or I'll go there. So if the probability was 1 that I was in this cell, then at the next level, it'll be half and a half. And at the next level, those halves will split again.

So the probability that I'm in a particular cell at level i is this P_i . And the probability that I'm in a

particular cell level i plus 1 is this P_i plus one. And there's some sparse matrix A which spreads that probability out. It splits it into my neighbors 50/50.

Here's a simulation of Plinko. So I started with the probability 1 in the center cell. And as I go through different levels, I get split 50/50. And you see a binomial or almost Gaussian distribution spread as I go through more and more levels until it's equally probable that I could wind up in any one of the cells.

You can think about it this way, right? The probability at level i plus 1 that the chip is in cell N is inherited 50/50 from its two neighbors, right? There's some probability that was in these two neighbors. I would inherit half of that probability. It would be split by these pegs.

The sparse matrix that represents this operation has two diagonals. And on each of those diagonals is a half. And you can build that matrix using the `spdiags` command.

It says that there's going to be two diagonal components which are equal to a half. And their position is going to be one on either side of the central diagonal. That's going to indicate that I pass this probability, 50/50, to each of my neighbors.

And then successive multiplications by A will split this probability. And we'll see the simulation that tells us how probable it is to find the Plinko chip in a particular column. Yes?

AUDIENCE: [INAUDIBLE]

JAMES W. SWAN: Yeah. So in diffusion in general?

AUDIENCE: Well, in this instance in particular because [INAUDIBLE]

JAMES W. SWAN: Well, OK. That's fair enough. This is one particular model of the Plinko board, which sort of imagines alternating cells that I'm falling through. We could construct an alternative model, if we wanted to, that didn't have that part of the picture, OK?

So that's a matrix that looks like this, right? The central diagonal is 0. Most of the off-diagonal components here are 0 and 1 above and 1 below. I get a half and a half.

And if I'm careful-- somebody mentioned I need boundary conditions. When the Plinko chip gets to the edge, it doesn't fall out of the game. It gets reflected back in. So maybe we have to choose some special values for a couple of elements of this matrix.

But this is a sparse matrix. It has a sparse structure. It models a diffusion problem, just like we saw before. Most of physics is local, like this, right? I just need to know what's going on with my neighbors. And I spread the probability out. I get this nice diffusion problem. So it looks like this.

Here's something to notice. After many levels or cycles, I multiply by A many, many times. This probability distribution always seems to flatten out. It becomes uniform.

It turns out there are even special distributions for which A times A times that distribution is equal to that distribution. You can see it at the end here. This is one of those special distributions where the probability is equal in every other cell, right?

And at the next level, it all gets passed down. That's one multiplication by-- it all gets spread by 50%. And the next multiplication, everything gets spread by 50% again. And I recover the same distribution that I had before, this uniform distribution.

That's a special distribution for which A times A times P is equal to P . And this distribution is one of the eigenvectors of this matrix A times A . It's a particular vector that when I multiply it by this matrix AA , I get that vector back.

It happens to be unstretched. So this vector points in some direction. I transform it by the matrix. And I get back something that points in the same direction. That's the definition of this thing called an eigenvector. And this will be the subject that we focus on today.

So eigenvectors of a matrix-- they're special vectors that are stretched on multiplication by the matrix. So they're transformed. But they're only transformed into a stretched form of whatever they were before. They point in a direction. You transform them by the matrix. And you get something that points in the same direction, but is stretched.

Before, we saw the amount of stretch. The previous example, we saw the amount of stretch was 1. It wasn't stretched at all. You just get back the same vector you had before. But in principle, it could come back with any length.

For a real N -by- N matrix, there will be eigenvectors and eigenvalues, which are the amount of stretch, which are complex numbers. And finding eigenvector-eigenvalue pairs involves solving N equations. We'd like to know what these eigenvectors and eigenvalues are.

They're non-linear because they depend on both the value and the vector, the product of the

two, for $N + 1$ unknowns. We don't know how to solve non-linear equations yet. So we're kind of-- might seem like we're in a rough spot. But I'll show you that we're not.

But because there's N equations for $N + 1$ unknowns, that means eigenvectors are not unique. If W is an eigenvector, then any other vector that points in that same direction is also an eigenvector, right? It also gets stretched by this factor λ .

So we can never say what an eigenvector is uniquely. We can only prescribe its direction. Whatever its magnitude is, we don't care. We just care about its direction.

The amount of stretch, however, is unique. It's associated with that direction. So you have an amount of stretch. And you have a direction. And that describes the eigenvector-eigenvalue pair.

Is this clear? You've heard of eigenvalues and eigenvectors before? Good.

So how do you find eigenvalues? They seem like special sorts of solutions associated with a matrix. And if we understood them, then we can do a transformation. So I'll explain that in a minute. But how do you actually find these things, these eigenvalues?

Well, I've got to solve an equation A times w equals λ times w , which can be transformed into A minus λ identity times w equals 0 . And so the solution set to this equation is either w is equal to 0 .

That's one possible solution to this problem or the eigenvector w belongs to the null space of this matrix. It's one of those special vectors that when it multiplies this matrix gives back 0 , right? It gets projected out on transformation by this matrix.

Well, this solution doesn't seem very useful to us, right? It's trivial. So let's go with this idea that w belongs to the null space of A minus λI . That means A minus λI must be a singular matrix, whatever it is, right? And if it's singular, then the determinant of A minus λI must be equal to 0 .

So if this is true, and it should be true if we don't want a trivial solution, then the determinant of A minus λI is equal to 0 . So if we can compute that determinant and solve for λ , then we'll know the eigenvalue.

Well, it turns out that the determinant of a matrix like A minus λI is a polynomial in terms

of λ . It's a polynomial of degree N called the characteristic polynomial. And the N roots of this characteristic polynomial are called the eigenvalues of the matrix.

So there are N possible λ s for which $A - \lambda I$ become singular. It has a null space. And associated with those values are eigenvectors, vectors that live in that null space.

So this polynomial-- we could compute it for any matrix. We could compute this thing in principle, right? And we might even be able to factor it into this form. And then λ_1 , λ_2 , λ_N in this factorized form are all the possible eigenvalues associated with our matrix A , right?

There are all the possible amounts of stretch that can be imparted to particular eigenvectors. We don't know those vectors yet, right? We'll find them in a second. But we know the amounts of stretch that can be imparted by this matrix. OK? Any questions so far? No. Let's do an example.

Here's a matrix, minus 2, 1, 3. And it's 0's everywhere else. And we'd like to find the eigenvalues of this matrix. So we need to know $A - \lambda I$ and its determinant. So here's $A - \lambda I$. We just subtract λ from each of the diagonals.

And the determinant-- well, here, it's just the product of the diagonal elements. So that's the determinant of a diagonal matrix like this, the product of the diagonal elements. So it's minus 2 minus λ times 1 minus λ times 3 minus λ .

And the determinant of this has to be equal to 0. So the amounts of stretch, the eigenvalues imparted by this matrix, are minus 2, 1, and 3. And we found the eigenvalues.

Here's another matrix. Can you work out the eigenvalues of this matrix? Let's take 90 seconds. You can work with your neighbors. See if you can figure out the eigenvalues of that matrix. Nobody's collaborating today. I'm going to do it myself.

AUDIENCE: [INAUDIBLE]

JAMES W. SWAN: It's OK.

OK. What are you finding? Anyone want to guess what are the eigenvalues?

AUDIENCE: [INAUDIBLE]

JAMES W. SWAN: Good. OK. So we need to compute the determinant of $A - \lambda I$. That'll be $(-2 - \lambda)(-2 - \lambda - 1)$.

You can solve this to find that λ equals -3 or -1 . These little checks are useful. If you couldn't do this, that's OK. But you should try to practice this on your own to make sure you can.

Here are some more examples. So the elements of a diagonal matrix are always the eigenvalues because the determinant of a diagonal matrix is the product of the diagonal elements. So these diagonal values here are the roots of the secular characteristic polynomial. They are the eigenvalues.

It turns out the diagonal elements of a triangular matrix are eigenvalues, too. This should seem familiar to you. We talked about easy-to-solve systems of equations, right? Diagonal systems of equations are easy to solve, right? Triangular systems of equations are easy to solve. It's also easy to find their eigenvalues.

So the diagonal elements here are the eigenvalues of the triangular matrix. And eigenvalues have certain properties that can be inferred from the properties of polynomials, right? Since they are the roots to a polynomial, if we know certain things that should be true of those polynomial roots, that has to be true of the eigenvalues themselves.

So if we have a matrix which is real-valued, then we know that we're going to have this polynomial of degree N which is also real-valued, OK? It can have no more than N roots, right? And so A can have no more than N distinct eigenvalues.

The eigenvalues, like the factors of the polynomial, don't have to be distinct, though? You could have multiplicity in the roots of the polynomial.

So it's possible that $\lambda = 1$ here is an eigenvalue twice. That's referred to as algebraic multiplicity. We'll come back to that idea in a second.

Because the polynomial is real-valued, it means that the eigenvalues could be real or complex, just like the roots of a real-valued polynomial. But complex eigenvalues always appear as conjugate pairs. If there is a complex eigenvalue, then necessarily its complex conjugate is also an eigenvalue.

And here's a couple other properties. So the determinant of a matrix is the product of the

eigenvalues. We talked once about the trace of a matrix, which is the sum of its diagonal elements. The trace of a matrix is also the sum of the eigenvalues. These can sometimes come in handy-- not often, but sometimes.

Here's an example I talked about before-- so a series of chemical reactions. So we have a batch, a batch reactor. We load some material in. And we want to know how the concentrations of A, B, C, and D vary as a function of time.

And so A transforms into B. B and C are in equilibrium. C and D are in equilibrium. And our conservation equation for material is here.

This is a rate matrix. We'd like to understand what the characteristic polynomial of that is. The eigenvalues of that matrix are going to tell us something about how different rate processes evolve in time.

You can imagine just using units. On this side, we have concentration over time. On this side, we have concentration. And the rate matrix has units of rate, or 1 over time. So those eigenvalues also have units of rate. And they tell us the rate at which different transformations between these materials occur.

And so if we want to find the characteristic polynomial of this matrix and we need to compute the determinant of this matrix minus λI -- so subtract λ from each of the diagonals-- even though this is a four-by-four matrix, its determinant is easy to compute because it's full of zeros. I'm not going to compute it for you here.

It'll turn out that the characteristic polynomial looks like this. You should actually try to do this determinant and show that the polynomial works out to be this. But knowing that this is the characteristic polynomial, what are the eigenvalues of the rate matrix? If that's the characteristic polynomial, what are the eigenvalues, or tell me some of the eigenvalues of the rate matrix?

AUDIENCE: 0.

JAMES W. SWAN: 0. 0's an eigenvalue. λ equals 0 is a solution. Minus k_1 is another solution. What is this eigenvalue 0 correspond to? What's that?

AUDIENCE: [INAUDIBLE]

JAMES W. SWAN: OK. Physically, it's a rate process with 0 rate, steady state. So the 0 eigenvalue's going to correspond to the steady state. The eigenvector associated with that eigenvalue should correspond to the steady state solution.

How about this eigenvalue minus k_1 ? This is a rate process with rate k_1 . What physical process does that represent? It's something evolving in time now, right?

So that's the transformation of A into B. And the eigenvector should reflect that transformation. We'll see what those eigenvectors are in a minute. But these eigenvalues can be interpreted in terms of physical processes.

This quadratic solution here has some eigenvalue. I don't know what it is. You use the quadratic formula and you can find it. But it involves k_2 , k_3 , k_4 . And this is a typo. It should be k_5 .

And so that says something about the interconversion between B, C, and D, and the rate processes that occur as we convert from B to C to D.

Is that too fast? Do you want to write some more on this slide before I go on, or are you OK? Are there any questions about this? No.

Given an eigenvalue, a particular eigenvalue, what's the corresponding eigenvector? We know the eigenvector isn't uniquely specified. It belongs to the null space of this matrix A minus λI times identity.

Even though it's not unique, we might still try to find it using Gaussian elimination, right? So we may try to take-- we may try to solve the equation A minus λI times identity multiplied by w equals 0 using Gaussian elimination.

But because it's not unique, at some point, we'll run out of rows to eliminate, right? There's a null space to this matrix, right? We won't be able to eliminate everything. We'd say it's rank deficient, right?

So we'll be able to eliminate up to some R, the rank of this matrix. And then all the components below are essentially free or arbitrarily specified. There are no equations to say what those components of the eigenvector are.

The number of all 0 rows-- it's called the geometric multiplicity of the eigenvalue. Sorry.

Geometric is missing here. It's the number of components of the eigenvector that can be freely specified.

The geometric multiplicity might be 1. That's like saying that the eigenvectors are all pointing in the same direction, but can have arbitrary magnitude, right?

It might have geometric multiplicity 2, which means the eigenvectors associated with this eigenvalue live in some plane. And any vector from that plane is a corresponding eigenvector. It might have a higher geometric multiplicity associated with it.

So let's try something here. Let's try to find the eigenvectors of this matrix. I told you what the eigenvalues were. They were the diagonal values here. So they're minus 2, 1, and 3.

Let's look for the eigenvector corresponding to this eigenvalue. So I want to solve this equation A minus this particular λ , which is minus 2, times identity equals 0. So I got to do Gaussian elimination on this matrix. It's already eliminated for me, right?

I have one row which is all 0's, which says the first component of my eigenvector can be freely specified. The other two components have to be 0. 3 times the second component of my eigenvector is 0. 5 times the third component is 0. So the other two components have to be 0. But the first component is freely specified.

So the eigenvector associated with this eigenvalue is 1, 0, 0. If I take a vector which points in the x-direction in R^3 and I multiply it by this matrix, it gets stretched by minus 2. So I point in the other direction. And I stretch out by a factor of 2.

You can guess then what the other eigenvectors are. What's the eigenvector associated with this eigenvalue here? 0, 1, 0, or anything proportional to that. What's the eigenvector associated with this eigenvalue? 0, 0, 1, or anything proportional to it.

All these eigenvectors have a geometric multiplicity of 1, right? I can just specify some scalar variant on them. And they'll transform into themselves.

Here's a problem you can try. Here's our series of chemical reactions again. And we want to know the eigenvector of the rate matrix having eigenvalue 0. This should correspond to the steady state solution of our ordinary differential equation here.

So you've got to do elimination on this matrix. Can you do that? Can you find this eigenvector?

Try it out with your neighbor. See if you can do it. And then we'll compare results. This will just be a quick test of understanding.

Are you guys able to do this? Sort of, maybe? Here's the answer, or an answer, for the eigenvector. It's not unique, right? It's got some constant out in front of it.

So you do Gaussian elimination here. So subtract or add the first row to the second row. You'll eliminate this 0, right? And then add the second row to the third row. You'll eliminate this k_2 .

You have to do a little bit more work to do elimination of k_4 here. But that's not a big deal. Again, you'll add the third row to the fourth row and eliminate that. And you'll also wind up eliminating this k_5 . So the last row here will be all 0's.

And that means the last component of our eigenvector's freely specifiable. It can be anything we want. So I said it is 1. And then I did back substitution to determine all the other components, right? That's the way to do this.

And here's what the eigenvector looks like when you're done. The steady state solution has no A in it. Of course, A is just eliminated by a forward reaction. So if we let this run out to infinity, there should be no A. And that's what happens.

But there's equilibria between B, C, and D. And the steady state solution reflects that equilibria. We have to pick what this constant out in front is. And we discussed this before, actually, right? You would pick that based on how much material was initially in the reactor.

We've got to have an overall mass balance. And that's missing from this system of equations, right? Mass conservation is what gave the null space for this rate matrix in the first place. Make sense?

Try this example out. See if you can work through the details of it. I think it's useful to be able to do these sorts of things quickly. Here are some simpler problems.

So here's a matrix. It's not a very good matrix. Matrices can't be good or bad. It's not particularly interesting. But it's all 0's. So what are its eigenvalues? It's just 0, right? The diagonal elements are the eigenvalues. And they're 0.

That eigenvalue has algebraic multiplicity 2. It's a double root of the secular characteristic polynomial. Can you give me the eigenvectors? Can you give me eigenvectors of this matrix?

Can you give me linearly independent-- yeah?

AUDIENCE: [INAUDIBLE]

JAMES W. SWAN: OK.

AUDIENCE: [INAUDIBLE]

JAMES W. SWAN: OK. Good. So this is a very ambiguous sort of problem or question, right? Any vector I multiply by A here is going to be stretched by 0 because A by its very nature is all 0's. All those vectors live in a plane. So any vector from that plane is going to be transformed in this way.

The eigenvector corresponding to eigenvalue 0 has geometric multiplicity 2 because I can freely specify two of its components. Oh my goodness. I went so fast. We'll just do it this way.

Algebraic multiplicity 2, geometric multiplicity 2-- I can pick two vectors. They can be any two I want in principle, right? It has geometric multiplicity 2.

Here's another matrix. It's a little more interesting than the last one. I stuck a 1 in there instead. Again, the eigenvalues are 0. It's a double root. So it has algebraic multiplicity 2.

But you can convince yourself that there's only one direction that transforms that squeeze down to 0, right? There's only one vector direction that lives in the null space of A minus λI -- lives in the null space of A . And that's vectors parallel to $(1, 0)$. So the eigenvector associated with that eigenvalue 0 has geometric multiplicity 1 instead of geometric multiplicity 2.

Now, here's an example for you to do. Can you find the eigenvalues and some linearly independent eigenvectors of this matrix, which looks like the one we just looked at. But now it's three-by-three instead of two-by-two. And if you find those eigenvalues and eigenvectors, what are the algebraic and geometric multiplicity?

Well, you guys must had a rough week. You're usually much more talkative and energetic than this.

[LAUGHTER]

Well, what are the eigenvalues here?

AUDIENCE: 0.

JAMES W. SWAN: Yeah. They all turn out to be 0. So that's an algebraic multiplicity of 3. It'll turn out there are two vectors, two vector directions, that I can specify that will both be squeezed down to 0. In fact, any vector from the x-y plane will also be squeezed down to 0. So this has algebraic multiplicity 3 and geometric multiplicity 2.

I'm going to explain why this is important in a second. But understanding that this can happen is going to be useful for you. So if an eigenvalue is distinct, then it has algebraic multiplicity 1. It's the only eigenvalue with that value. It's the only time that amount of stretch is imparted.

And there will be only one corresponding eigenvector. There will be a direction and an amount of stretch. If an eigenvalue has an algebraic multiplicity M , well, you just saw that the geometric multiplicity, which is the dimension of the null space of $A - \lambda I$ -- it's the dimension of the space spanned by no vectors of $A - \lambda I$ -- it's going to be bigger than 1 or equal to 1. And it's going to be smaller or equal to M .

And we saw different variants on values that sit in this range. So there could be as many as M linearly independent eigenvectors. And there may be fewer.

So geometric multiplicity-- it's the number of linearly independent eigenvectors associated with an eigenvalue. It's the dimension of the null space of this matrix.

Problems for which the geometric and algebraic multiplicity are the same for all the eigenvalues and eigenvectors, all those pairs, are nice because the matrix then is said to have a complete set of eigenvectors. There's enough eigenvectors in the problem that they describe the span of our vector space \mathbb{R}^N that our matrix is doing transformations between.

If we have geometric multiplicity that's smaller than the algebraic multiplicity, then some of these stretched-- we can't stretch in all possible directions in \mathbb{R}^N . There's going to be a direction that might be left out.

We want to be able to do a type of transformation called an eigendecomposition. I'm going to show you that in a second. It's useful for solving systems of equations or for transforming systems of ordinary differential equations, linear ordinary differential equations.

But we're only going to be able to do that when we have this complete set of eigenvectors. When we don't have that complete set, we're going to have to do other sorts of

transformations.

You have a problem in your homework now, I think, that has this sort of a hang-up associated with it. It's the second problem in your homework set. That's something to think about.

For a matrix with the complete set of eigenvectors, we can write the following. A times a matrix W is equal to W times the matrix λ . Let me tell you what W and λ are.

So W 's a matrix whose columns are made up of this-- all of these eigenvectors. And λ 's a matrix whose diagonal values are each of the corresponding eigenvalues associated with those eigenvectors.

This is nothing more than a restatement of the original eigenvalue problem. AW is λW . But now each eigenvalue has a corresponding particular eigenvector. And we've stacked those equations up to make this statement about matrix-matrix multiplication.

So we've taken each of these W 's over here. And we've just made them the columns of a particular matrix. But it's nothing more than a restatement of the fundamental eigenvalue problem we posed at the beginning here.

But what's nice is if I have this complete set of eigenvectors, then W has an inverse that I can write down. So another way to state this same equation is that λ -- the eigenvalues can be found from this matrix product, $W^{-1}AW$.

And under these circumstances, we say the matrix can be diagonalized. There's a transformation from A to a diagonal form. That's good for us, right? We know diagonal systems of equations are easy to solve, right?

So if I knew what the eigenvectors were, then I can transform my equation to this diagonal form. I could solve systems of equations really easily.

Of course, we just saw that knowing what those eigenvectors are requires solving systems of equations, anyway. So the problem of finding the eigenvectors is as hard as the problem of solving a system of equations. But in principle, I can do this sort of transformation.

Equivalently, the matrix A can be written as $W\lambda W^{-1}$. These are all equivalent ways of writing this fundamental relationship up here when the inverse of W exists.

So this means that if I know the eigenvalues and eigenvectors, I can easily reconstruct my

equation, right? If I know the eigenvectors in A, then I can easily diagonalize my system of equations, right? So this is a useful sort of transformation to do.

We haven't talked about how it's done in the computer. We've talked about how you would do it by hand. These are ways you could do it by hand. The computer won't do Gaussian elimination for each of those eigenvectors independently, right?

Each elimination procedure is order N cubed, right? And you got to do that for N eigenvectors. So that's N to the fourth operations. That's pretty slow.

There's an alternative way of doing it that's beyond the scope of this class called-- it's called the Lanczos algorithm. And it's what's referred to as a Krylov subspace method, that sort of iterative method where you take products of your matrix with certain vectors and from those products, infer what the eigenvectors and eigenvalues are.

So that's the way a computer's going to do it. That's going to be an order N cubed sort of calculation to find all the eigenvalues and eigenvectors [INAUDIBLE] solving a system of equations. But sometimes you want these things.

Here's an example of how this eigendecomposition can be useful to you if you did it. So we know the matrix A can be represented as $W \lambda W^{-1} x = b$. This is our transformed system of equations here. We've just substituted for A.

If I multiply both sides of this equation by W^{-1} , then I've got λ times the quantity $W^{-1} x$ is equal to $W^{-1} b$. And if I call this quantity in parentheses y , then I have an easy-to-solve system of equations for y .

y is equal to $\lambda^{-1} c$. But λ^{-1} is just 1 over each of the diagonal components of λ . λ 's a diagonal matrix.

Then all I need to do-- ooh, typo. There's an equal sign missing here. Sorry for that. Now all I need to do is substitute for what I called y and what I called c .

So y was $W^{-1} x$. That's equal to $\lambda^{-1} W^{-1} b$. And so I multiply both sides of this equation by W . And I get $x = W \lambda^{-1} W^{-1} b$.

So if I knew the eigenvalues and eigenvectors, I can really easily solve the system of equations. If I did this decomposition, I could solve many systems of equations, right? They're

simple to solve with just matrix-matrix multiplication.

Now, how is W inverse computed? Well, W inverse transpose are actually the eigenvectors of A transpose. You may have to compute this matrix explicitly. But there are times when we deal with so-called symmetric matrices, ones for which they are equal to their transpose.

And if that's the case, and if you take all of your eigenvectors and you normalize them so they're of length 1-- the Euclidean norm is 1-- then it'll turn out that W inverse is precisely equal to W transpose, right? And so the eigenvalue matrix will be unitary. It'll have this property where its transposes is its inverse, right?

So this becomes trivial to do then, this process of W inverse. It's not always true that this is the case, right? It is true when we deal with problems that have symmetric matrices associated with them. That pops up in a lot of cases.

You can prove-- I might ask you to show this some time-- that the eigenvectors of a symmetric matrix are orthogonal, that they satisfy this property that-- I take the dot product between two different eigenvectors and it'll be equal to 0 unless those are the same eigenvector. That's a property associated with symmetric matrices.

They're also useful when analyzing systems of ordinary differential equations. So here, I've got a differential equation, a vector x dot. So the time derivative of x is equal to A times x .

And if I substitute my eigendecomposition-- so $W \lambda W$ inverse-- and I define a new unknown y instead of x , then I can diagonalize that system of equations. So you see y dot is equal to λ times y where each component of y is decoupled from all of the others. Each of them satisfies their own ordinary differential equation that's not coupled to any of the others, right? And it has a simple first-order rate constant, which is the eigenvalue associated with that particular eigendirection.

So this system of ODEs is decoupled. And it's easy to solve. You know the solution, right? It's an exponential. And that can be quite handy when we're looking at different sorts of chemical rate processes that correspond to linear differential equations.

We'll talk about nonlinear, systems of nonlinear, differential equations later in this term. And you'll find out that this same sort of analysis can be quite useful there. So we'll linearize those equations. And we'll ask is their linear-- in their linearized form, what are these different rate constants? How big are they?

They might determine what we need to do in order to integrate those equations numerically because there are many times when there's not a complete set of eigenvectors. That happens. And then the matrix can't be diagonalized in this way.

There are some components that can't be decoupled from each other. That's what this diagonalization does, right? It splits up these different stretching directions from each other. But there's some directions that can't be decoupled from each other anymore.

And then there are other transformations one can do. So there's an almost diagonal form that you can transform into called the Jordan normal form.

There are other transformations that one can do, like called, for example, Schur decomposition, which is a transformation into an upper triangular form for this matrix. We'll talk next time about the singular value decomposition, which is another sort of transformation one can do when we don't have these complete sets of eigenvectors.

But this concludes our discussion of eigenvalues and eigenvectors. You'll get a chance to practice these things on your next two homework assignments, actually. So it'll come up in a couple of different circumstances.

I would really encourage you to try to solve some of these example problems that were in here. Solving by hand can be useful. Make sure you can work through the steps and understand where these different concepts come into play in terms of determining what the eigenvalues and eigenvectors are.

All right. Have a great weekend. See you on Monday.