Yes, OK, four, three, two, one, OK, I see you guys are in a happy mood.

I don't know if that means 18.06 is ending, or, the quiz was good. Uh, my birthday conference was going on at the time of the quiz, and in the conference, of course, everybody had to say nice things, but I was wondering, what would my 18.06 class be saying, because it was at the exactly the same time.

But, what I know from the grades so far, they're basically close to, and maybe slightly above the grades that you got on quiz two. So, very satisfactory.

And, then we have a final exam coming up, and today's lecture, as I told you by email, will be a first step in the review, and then on Wednesday I'll do all I can in reviewing the whole course. So my topic today is -- actually, this is a lecture I have never given before in this way, and it will -- well, four subspaces, that's certainly fundamental, and you know that, so I want to speak about left-inverses and right-inverses and then something called pseudo-inverses.

And pseudo-inverses, let me say right away, that comes in near the end of chapter seven, and that would not be expected on the final.

But you'll see that what I'm talking about is really the basic stuff that, for an m-by-n matrix of rank r, we're going back to the most fundamental picture in linear algebra. Nobody could forget that picture, right? When you're my age, even, you'll remember the row space, and the null space.

Orthogonal complements over there, the column space and the null space of A transpose column, orthogonal complements over here. And I want to speak about inverses. OK.

And I want to identify the different possibilities.

So first of all, when does a matrix have a just a perfect inverse, two-sided, you know, so the two-sided inverse is what we just call inverse, right? And, so that means that there's a matrix that produces the identity, whether we write it on the left or on the right. And just tell me, how are the numbers r, the rank, n the number of columns, m the number of rows, how are those numbers related when we have an invertible matrix? So this is the matrix which was -- chapter two was all about matrices like this, the beginning of the course, what was the relation of th- of r, m, and n, for the nice case? They're all the same, all equal.

So this is the case when r=m=n. Square matrix, full rank, period, just -- so I'll use the words full rank. OK, good. Everybody knows that. OK.

Then chapter three. We began to deal with matrices that were not of full rank, and they could have any rank, and we learned what the rank was.

And then we focused, if you remember on some cases like full column rank. Now, can you remember what was the deal with full column rank? So, now, I think this is the case in which we have a left-inverse, and I'll try to find it. So we have a -- what was the situation there? It's the case of full column rank, and that means -- what does that mean about r? It equals, what's the deal with r, now, if we have full column rank, I mean the columns are independent, but maybe not the rows. So what is r equal to in this case? n.

Thanks. n.

$r=n$. The n columns are independent, but probably, we have more rows.

What's the picture, and then what's the null space for this? So the n columns are independent, what's the null space in this case? So of course, you know what I'm asking.

You're saying, why is this guy asking something, I know that-- I think about it in my sleep, right? So the null space of this matrix if the rank is n, the null space is what vectors are in the null space? Just the zero vector.

Right? The columns are independent.

Independent columns. No combination of the columns gives zero except that one. And what's my picture over, -- let me redraw my picture -- the row space is everything.

No. Is that right? Let's see, I often get these turned around, right? So what's the deal? The columns are independent, right? So the rank should be the full number of columns, so what does that tell us? There's no null space, right. OK. The row space is the whole thing.

Yes, I won't even draw the picture.

And what was the deal with -- and these were very important in least squares problems because -- So, what more is true here? If we have full column rank, the null space is zero, we have independent columns, the unique -- so we have zero or one solutions to Ax=b. There may not be any solutions, but if there's a solution, there's only one solution because other solutions are found by adding on stuff from the null space, and there's nobody there to add on. So the particular solution is the solution, if there is a particular solution. But of course, the rows might not be - are probably not independent -- and therefore, so right-hand sides won't end up with a zero equal zero after elimination, so sometimes we may have no solution, or one solution.

OK. And what I want to say is that for this matrix A -- oh, yes, tell me something about A transpose A in this case. So this whole part of the board, now, is devoted to this case.

What's the deal with A transpose A? I've emphasized over and over how important that combination is, for a rectangular matrix, A transpose A is the good thing to look at, and if the rank is n, if the null space has only zero in it, then the same is true of A transpose A.

That's the beautiful fact, that if the rank of A is n, well, we know this will be an n by n symmetric matrix, and it will be full rank. So this is invertible.

This matrix is invertible. That matrix is invertible.

And now I want to show you that A itself has a one-sided inverse. Here it is.

The inverse of that, which exists, times A transpose, there is a one-sided -- shall I call it A inverse? -- left of the matrix A.

Why do I say that? Because if I multiply this guy by A, what do I get? What does that multiplication give? Of course, you know it instantly, because I just put the parentheses there, I have A transpose A inverse times A transpose A so, of course, it's the identity. So it's a left inverse.

And this was the totally crucial case for least squares, because you remember that least squares, the central equation of least squares had this matrix, A transpose A, as its coefficient matrix. And in the case of full column rank, that matrix is invertible, and we're go.

So that's the case where there is a left-inverse. So A does whatever it does, we can find a matrix that brings it back to the identity. Now, is it true that, in the other order -- so A inverse left times A is the identity. Right? This matrix is m by n. This matrix is n by m. The identity matrix is n by n. All good.

All good if you're n. But if you try to put that matrix on the other side, it would fail.

If the full column rank -- if this is smaller than m, the case where they're equals is the beautiful case, but that's all set. Now, we're looking at the case where the columns are independent but the rows are not. So this is invertible, but what matrix is not invertible? A A transpose is bad for this case. A transpose A is good. So we can multiply on the left, everything good, we get the left inverse.

But it would not be a two-sided inverse.

A rectangular matrix can't have a two-sided inverse, because there's got to be some null space, right? If I have a matrix that's rectangular, then either that matrix or its transpose has some null space, because if n and m are different, then there's going to be some free variables around, and we'll have some null space in that direction. OK, tell me the corresponding picture for the opposite case. So now I'm going to ask you about right-inverses. A right-inverse.

And you can fill this all out, this is going to be the case of full row rank. And then r is equal to m, now, the m rows are independent, but the columns are not. So what's the deal on that? Well, just exactly the flip of this one.

The null space of A transpose contains only zero, because there are no combinations of the rows that give the zero row. We have independent rows. And in a minute, I'll give an example of all these. So, how many solutions to Ax=b in this case? The rows are independent. So we can always solve Ax=b. Whenever elimination never produces a zero row, so we never get into that zero equal one problem, so Ax=b always has a solution, but too many. So there will be some null space, the null space

of A -- what will be the dimension of A's null space? How many free variables have we got? How many special solutions in that null space have we got? So how many free variables in this setup? We've got n columns, so n variables, and this tells us how many are pivot variables, that tells us how many pivots there are, so there are n-m free variables.

So there are infinitely many solutions to Ax=b.

We have n-m free variables in this case.

OK. Now I wanted to ask about this idea of a right-inverse. OK. So I'm going to have a matrix A, my matrix A, and now there's going to be some inverse on the right that will give the identity matrix. So it will be A times A inverse on the right, will be I.

And can you tell me what, just by comparing with what we had up there, what will be the right-inverse, we even have a formula for it. There will be other -- actually, there are other left-inverses, that's our favorite. There will be other right-inverses, but tell me our favorite here, what's the nice right-inverse? The nice right-inverse will be, well, there we had A transpose A was good, now it will be A A transpose that's good. The good matrix, the good right -- the thing we can invert is A A transpose, so now if I just do it that way, there sits the right-inverse. You see how completely parallel it is to the one above? Right.

So that's the right-inverse. So that's the case when there is -- In terms of this picture, tell me what the null spaces are like so far for these three cases. What about case one, where we had a two-sided inverse, full rank, everything great.

The null spaces were, like, gone, right? The null spaces were just the zero vectors.

Then I took case two, this null space was gone.

Case three, this null space was gone, and then case four is, like, the most general case when this picture is all there -- when all the null spaces -- this has dimension r, of course, this has dimension n-r, this has dimension r, this has dimension m-r, and the final case will be when r is smaller than m and n. But can I just, before I leave here look a little more at this one? At this case of full column rank? So A inverse on the left, it has this left-inverse to give the identity. I said if we multiply it in the other order, we wouldn't get the identity.

But then I just realized that I should ask you, what do we get? So if I put them in the other order -- if I continue this down below, but I write A times A inverse left -- so there's A times the left-inverse, but it's not on the left any more.

So it's not going to come out perfectly. But everybody in this room ought to recognize that matrix, right? Let's see, is that the guy we know? Am I OK, here? What is that matrix? P.

Thanks. P.

That matrix -- it's a projection.

It's the projection onto the column space.

It's trying to be the identity matrix, right? A projection matrix tries to be the identity matrix, but you've given it, an impossible job.

So it's the identity matrix where it can be, and elsewhere, it's the zero matrix.

So this is P, right.

A projection onto the column space. OK. And if I asked you this one, and put these in the opposite order -- so this came from up here. And similarly, if I try to put the right inverse on the left -- so that, like, came from above. This, coming from this side, what happens if I try to put the right inverse on the left? Then I would have A transpose A, A transpose inverse A, if this matrix is now on the left, what do you figure that matrix is? It's going to be a projection, too, right? It looks very much like this guy, except the only difference is, A and A transpose have been reversed. So this is a projection, this is another projection, onto the row space.

Again, it's trying to be the identity, but there's only so much the matrix can do. And this is the projection onto the column space. So let me now go back to the main picture and tell you about the general case, the pseudo-inverse. These are cases we know.

So this was important review. You've got to know the business about these ranks, and the free variables -- really, this is linear algebra coming together.

And, you know, one nice thing about teaching 18.06, It's not trivial.

But it's -- I don't know, somehow, it's nice when it comes out right. I mean -- well, I shouldn't say anything bad about calculus, but I will. I mean, like, you know, you have formulas for surface area, and other awful things and, you know, they do their best in calculus, but it's not elegant. And, linear algebra just is -- well, you know, linear algebra is about the nice part of calculus, where everything's, like, flat, and, the formulas come out right. And you can go into high dimensions where, in calculus, you're trying to visualize these things, well, two or three dimensions is kind of the limit. But here, we don't -- you know, I've stopped doing two-by-twos, I'm just talking about the general case. OK, now I really will speak about the general case here. What could be the inverse -- what's a kind of reasonable inverse for a matrix for the completely general matrix where there's a rank r, but it's smaller than n, so there's some null space left, and it's smaller than m, so a transpose has some null space, and it's those null spaces that are screwing up inverses, right? Because if a matrix takes a vector to zero, well, there's no way an inverse can, like, bring it back to life.

My topic is now the pseudo-inverse, and let's just by a picture, see what's the best inverse we could have? So, here's a vector x in the row space. I multiply by A.

Now, the one thing everybody knows is you take a vector, you multiply by A, and you get an output, and where is that output? Where is Ax? Always in the column space, right? Ax is a combination of the columns.

So Ax is somewhere here. So I could take all the vectors in the row space. I could multiply them all by A.

I would get a bunch of vectors in the column space and what I think is, I'd get all the vectors in the column space just right. I think that this connection between an x in the row space and an Ax in the column space, this is one-to-one. We got a chance, because they have the same dimension.

That's an r-dimensional space, and that's an r-dimensional space. And somehow, the matrix A -- it's got these null spaces hanging around, where it's knocking vectors to zero.

And then it's got all the vectors in between, which is almost all vectors. Almost all vectors have a row space component and a null space component.

And it's killing the null space component.

But if I look at the vectors that are in the row space, with no null space component, just in the row space, then they all go into the column space, so if I put another vector, let's say, y, in the row space, I positive that wherever Ay is, it won't hit Ax. Do you see what I'm saying? Let's see why. All right.

So here's what I said. If x and y are in the row space, then A x is not the same as A y.

They're both in the column space, of course, but they're different. That would be a perfect question on a final exam, because that's what I'm teaching you in that material of chapter three and chapter four, especially chapter three. If x and y are in the row space, then Ax is different from Ay.

So what this means -- and we'll see why -- is that, in words, from the row space to the column space, A is perfect, it's an invertible matrix.

If we, like, limited it to those spaces.

And then, its inverse will be what I'll call the pseudo-inverse. So that's that the pseudo-inverse is. It's the inverse -- so A goes this way, from x to y -- sorry, x to A x, from y to A y, that's A, going that way. Then in the other direction, anything in the column space comes from somebody in the row space, and the reverse there is what I'll call the pseudo-inverse, and the accepted notation is A plus. So y will be A plus x.

I'm sorry. No, y will be A plus times whatever it started with, A y.

Do you see my picture there? Same, of course, for x and A x. This way, A does it, the other way is the pseudo-inverse, and the pseudo-inverse just kills this stuff, and the matrix just kills this stuff.

So everything that's really serious here is going on in the row space and the column space, and now, tell me -- this is the fundamental fact, that between those two r-dimensional spaces, our matrix is perfect.

Why? Suppose they weren't.

Why do I get into trouble? Suppose -- so, proof. I haven't written down proof very much, but I'm going to use that word once.

Suppose they were the same. Suppose these are supposed to be two different vectors. Maybe I'd better make the statement correctly. If x and y are different vectors in the row space -- maybe I'll better put if x is different from y, both in the row space -- so I'm starting with two different vectors in the row space, I'm multiplying by A -- so these guys are in the column space, everybody knows that, and the point is, they're different over there. So, suppose they weren't.

Suppose $A x=A y$. Suppose, well, that's the same as saying $A(x-y)$ is zero.

So what? So, what do I know now about $(x-y)$, what do I know about this vector? Well, I can see right away, what space is it in? It's sitting in the null space, right? So it's in the null space. But what else do I know about it? Here it was x in the row space, y in the row space, what about x-y? It's also in the row space, right? Heck, that thing is a vector space, and if the vector space is anything at all, if x is in the row space, and y is in the row space, then the difference is also, so it's also in the row space. So what? Now I've got a vector x-y that's in the null space, and that's also in the row space, so what vector is it? It's the zero vector. So I would conclude from that that x-y had to be the zero vector, x-y, so, in other words, if I start from two different vectors, I get two different vectors.

If these vectors are the same, then those vectors had to be the same. That's like the algebra proof, which we understand completely because we really understand these subspaces of what I said in words, that a matrix A is really a nice, invertible mapping from row space to columns pace. If the null spaces keep out of the way, then we have an inverse.

And that inverse is called the pseudo inverse, and it's a very, very, useful in application. Statisticians discovered, oh boy, this is the thing that we needed all our lives, and here it finally showed up, the pseudo-inverse is the right thing. Why do statisticians need it? And because statisticians are like least-squares-happy. I mean they're always doing least squares. And so this is their central linear regression. Statisticians who may watch this on video, please forgive that description of your interests. One of your interests is linear regression and this problem. But this problem is only OK provided we have full column rank.

And statisticians have to worry all the time about, oh, God, maybe we just repeated an experiment.

You know, you're taking all these measurements, maybe you just repeat them a few times. You know, maybe they're not independent.

Well, in that case, that A transpose A matrix that they depend on becomes singular. So then that's when they needed the pseudo-inverse, it just arrived at the right moment, and it's the right quantity.

OK. So now that you know what the pseudo-inverse should do, let me see what it is.

Can we find it? So this is my -- to complete the lecture is -- how do I find this pseudo-inverse A plus? OK. OK.

Well, here's one way. Everything I do today is to try to review stuff. One way would be to start from the SVD. The Singular Value Decomposition. And you remember that

that factored A into an orthogonal matrix times this diagonal matrix times this orthogonal matrix.

But what did that diagonal guy look like? This diagonal guy, sigma, has some non-zeroes, and you remember, they came from A transpose A, and A A transpose, these are the good guys, and then some more zeroes, and all zeroes there, and all zeroes there. So you can guess what the pseudo-inverse is, I just invert stuff that's nice to invert -- well, what's the pseudo-inverse of this? That's what the problem comes down to. What's the pseudo-inverse of this beautiful diagonal matrix? But it's got a null space, right? What's the rank of this matrix? What's the rank of this diagonal matrix? r, of course. It's got r non-zeroes, and then it's otherwise, zip.

So it's got n columns, it's got m rows, and it's got rank r. It's the best example, the simplest example we could ever have of our general setup.

OK? So what's the pseudo-inverse? What's the matrix -- so I'll erase our columns, because right below it, I want to write the pseudo-inverse. OK, you can make a pretty darn good guess. If it was a proper diagonal matrix, invertible, if there weren't any zeroes down here, if it was sigma one to sigma n, then everybody knows what the inverse would be, the inverse would be one over sigma one, down to one over s- but of course, I'll have to stop at sigma r. And, it will be the rest, zeroes again, of course.

And now this one was m by n, and this one is meant to have a slightly different, you know, transpose shape, n by m. They both have that rank r.

My idea is that the pseudo-inverse is the best -- is the closest I can come to an inverse.

So what is sigma times its pseudo-inverse? Can you multiply sigma by its pseudo-inverse? Multiply that by that? What matrix do you get? They're diagonal. Rectangular, of course. But of course, we're going to get ones, R ones, and all the rest, zeroes. And the shape of that, this whole matrix will be m by m.

And suppose I did it in the other order. Suppose I did sigma plus sigma. Why don't I do it right underneath? in the opposite order? See, this matrix hasn't got a left-inverse, it hasn't got a right-inverse, but every matrix has got a pseudo-inverse. If I do it in the order sigma plus sigma, what do I get? Square matrix, this is m by n, this is m by m, my result is going to m by m -- is going to be n by n, and what is it? Those are diagonal matrices, it's going to be ones, and then zeroes. It's not the same as that, it's a different size -- it's a projection. One is a projection matrix onto the column space, and this one is the projection matrix onto the row space. That's the best that pseudo-inverse can do. So what the pseudo-inverse does is, if you multiply on the left, you don't get the identity, if you multiply on the right, you don't get the identity, what you get is the projection. It brings you into the two good spaces, the row space and column space.

And it just wipes out the null space.

So that's what the pseudo-inverse of this diagonal one is, and then the pseudo-inverse of A itself -- this is perfectly invertible. What's the inverse of V transpose? Just another tiny bit of review. That's an orthogonal matrix, and its inverse is V, good.

This guy has got all the trouble in it, all the null space is responsible for, so it doesn't have a true inverse, it has a pseudo-inverse, and then the inverse of U is U transpose, thanks. Or, of course, I could write U inverse. So, that's the question of, how do you find the pseudo-inverse -- so what statisticians do when they're in this -- so this is like the case of where least squares breaks down because the rank is -- you don't have full rank, and the beauty of the singular value decomposition is, it puts all the problems into this diagonal matrix where it's clear what to do.

It's the best inverse you could think of is clear.

You see there could be other -- I mean, we could put some stuff down here, it would multiply these zeroes.

It wouldn't have any effect, but then the good pseudo-inverse is the one with no extra stuff, it's sort of, like, as small as possible.

It has to have those to produce the ones.

If it had other stuff, it would just be a larger matrix, so this pseudo-inverse is kind of the minimal matrix that gives the best result. Sigma sigma plus being r ones.

SK. so I guess I'm hoping -- pseudo-inverse, again, let me repeat what I said at the very beginning. This pseudo-inverse, which appears at the end, which is in section seven point four, and probably I did more with it here than I did in the book. The word pseudo-inverse will not appear on an exam in this course, but I think if you see this all will appear, because this is all what the course was about, chapters one, two, three, four -- but if you see all that, then you probably see, well, OK, the general case had both null spaces around, and this is the natural thing to do. Yes.

So, this is one way to find the pseudo-inverse.

The point of a pseudo-inverse, of computing a pseudo-inverse is to get some factors where you can find the pseudo-inverse quickly. And this is, like, the champion, because this is where we can invert those, and those two, easily, just by transposing, and we know what to do with a diagonal. OK, that's as much review, maybe -- let's have a five-minute holiday in 18.06 and, I'll see you Wednesday, then, for the rest of this course.

Thanks.