[00:01:00] We're now recording.
All right, everybody, welcome to the inaugural webinar of this course.
Welcome to this course.
I apologise for the late start.
It looks like to those of you who are watching the video, apparently what happened is that Blackboard Collaborate was misconfigured, the link was incorrect, which resulted in everybody having to join via guest links, which caused the delays.
But in any case, I think we have a bit of time to spare at the moment.
Again, welcome to the course.
I don't know whether you've had a chance to watch the main introductory videos.
I would encourage you to do that and it sounds like everybody is pretty far along in that.
I'm actually going to focus on the logistics of this [00:02:00] course particularly since I think the way we're going to run some aspects here, I'm going to be a little bit different from other key path courses.
This is because experience has shown that at least for the maths and stats ones, the standard way you do things wasn't quite working out.
Nonetheless, let's hope this works out.
I'm mainly going to talk about the structure of the course and then I'll be happy to ask questions.
If you have a question, please click the raise hand item or type the question into chat.
Now, this is one common suggestion, having working with Blackboard Collaborate is that when you are typing a question into chat, I recommend you click on the raise hand first.
The reason for that is that by the time you're finished typing your question, I might have moved on, [00:03:00] and then we've lost perhaps some of the context for your question.
Again, press hand up, I will pause and wait for you to type your question, or you can unmute yourself and just ask the question verbally, that's fine too.
Any questions? We have one more thing actually if you click on the little status button.
Yes, Daniel.
Hello.
If you click on my status and settings button, you actually see some options like happy, sad, confused, surprised, faster, slower and agree or disagree.
I [00:04:00] will be using these a little bit.
Well, first of all, you're always free to do that to provide immediate feedback.
It is a little bit difficult to monitor for me, so I might miss it sometimes, but we'll see how well that will go.
Course learning outcomes, this course covers the area called multivariate analysis.
This area is defined a little bit traditionally.
It's sort of [inaudible 00:04:37].
A traditional multivariate analysis is basically methods that centre around the multivariate normal distribution.
I guess it's probably the best way to define it.
Obviously, we've moved beyond that so while a lot of what we are going to be working with is the multivariate normal distribution, which we'll introduce on week 1, we will, also [00:05:00] particularly later in the term, discuss other approaches which do not require it.
Nonetheless, when dealing with lots of variables, well, just if we're going to think of it mathematically, if you have p variables, that means you have p choose 2 pairwise interactions.
That is p times p -1/2, so quadratic and the number of pairwise interactions, and if we're going to consider all possible interactions, that's 2^p, so exponential, which means that multivariate models can get very complicated very quickly.
We require special methods and sometimes we require simplifying assumptions like multivariate

normality.
Now, the assessment will be structured as follows, we will be using Moodle and [00:06:00] also be using Turnitin for working just in case.
Each assessment, so there will be six quizzes, each weekly quiz, they should be pretty short.
The first one, well actually, I know there's some confusion and different little bits and documentation contradict each other.
The first quiz I will open up on Monday that is next Monday and that's what, in three days and they will not have a time limit.
Now, quizzes can consist of multiple choice or short answer questions or a code challenge.
Now in our code challenge, we'll talk about that later when I post the quiz.
Probably on webinar Monday, but the idea is that you will need to implement something in R.
Then you will be able to check your implementation against test cases [00:07:00] which will tell you if your code satisfies the requirements of the question.
Now the questions have a number of different forms as you will see throughout the course.
Then the final assessment will be after I have test your code against my own secret test cases to see if it handles situations beyond the ones you were given.
The remaining assessments, again, I'll talk about them closer to the actual assessment being posted.
Finally, there is the data analysis report.
I'll post information about it shortly, but basically you'll be given a substantively meaningful data analysis task.
Pretend you are a data scientist.
[00:08:00] Give me one sec I'll finish the sentence.
You pretend that you are a data scientist hired by a company to answer a particular question for them.
That is, for which methods are the scores are suitable.
Then I will role play the representative from your client's representative who will be able to answer questions, but that might not have all the statistical skills that you do.
Then you will need to submit an executive summary of your conclusions plus what's called an R Notebook with you're working.
All right, now the question.
Ben, I said the first quiz won't be timed.
Subsequent quizzes, I'm going to double-check, that I'm still sorting that out sorry.
[00:09:00] That's the analysis report.
It will be due at the end of Week 7 and you will receive individualised feedback for that.
Most of our interactions are going to be through these webinar sessions.
They're going to be three of them a week during regular weeks.
Because of my other scheduled commitments including other courses I'm teaching it, this is when I had to set them.
I tried to put them as late in the day as possible, because I know many of you have full-time jobs.
This is when they will be.
I will typically have some pre-planned material that I will discuss like now.
Then most of the rest of the time will be taken up answering questions asked on the forum.
Speaking of which just let me quickly check because somebody is posting something on the forum.
[00:10:00] Just by the way, in case anybody is monitoring the forum, the question was whether I'll be able to put it into recording after the session? The answer is yes.
The series sessions are recorded and this one you might have a little trouble accessing, but if not, I can download it and post it somewhere, so don't worry about that.
I will go through your questions and whenever I'm going to start on a question, I will make sure there's a timestamp that you can quickly see where it is in the recording.

This is mostly described here.
In order to keep things organised, I will be focusing my quick question answers in I the QA threads.
Outside of those threads, feel free to discuss things, but I reserve the right not to field questions from there.
Again, [00:11:00] the earlier keep path model put in practical in mathematics.
We're trying something new here.
That said, if you have a question about logistics such as when is this due, how much time do I have for this technical issues, like, I can't get into Blackboard, which is a very salient one at the moment.
Similar matters to address me as a course coordinator rather than the instructor, please post them in the forum, tag them appropriately and I will answer them.
Now, there was a question in the forum about reading lists.
Let me see if that was one of the ones in there.
Thank you for now, it's not.
Well, in the reading responses I'm going to enter any way this is Leganto reading lists.
You should have access to them from this link.
Let's see what that will do.
There it is.
[00:12:00] When you click that link, you should assume you're logged in, have access to these.
I think you might have access to some extracts from these books, if not, then unfortunately, this is locked down for you.
You might be able to find them somewhere else, I don't know.
But yes, this is the best we can do at the moment.
Those are optional readings, this is a bit about me.
I am a senior lecturer in statistics apparently.
I moved here in 2019 after six years of Wollongong.
I actually been around for quite awhile.
I actually come from the US and I've lived most of my life there and had most [00:13:00] of my education there.
Most of my research is in social networks or social network analysis and relational data in general, which you can view as a multivariate dependent categorical data, because you have relationships and those relationships can depend on each other.
A lot of variables computing.
I develop a bunch of popular R packages on CRAN, Comprehensive R Archive Network.
In fact, I just had to patch accepted into our core because there was a bug that I fixed, so it happens.
This is the basic structure.
Now, one item that I want to make very clear about the forum, because I think there's sometimes issues about that, [00:14:00] I will be posting announcements on the forum.
Some of those might be time sensitive, such as, there's an error in the quiz, here's the correction or hopefully that won't happen, but if in case it does, where various other clarification.
I would appreciate it if you checked your UNSW email at least once a day or every once every two days.
I know everybody is busy.
I know many of you have full-time jobs and other commitments, but it has been an issue before that I get emails about something that some email I sent on Monday.
I get an email the Monday after saying, well, I didn't know that this was announced, I will check my email on the weekend.
Again, [00:15:00] I'm sure everybody here will be very diligent, but please check announcements.
I will try not to flood your inbox, but please check announcements daily or every other day.
Here I'm going to pause for questions about the course logistics, if any.

[00:16:00] In that case, let's move on to a Q&A from the forum thread.
Now I've gotten three questions that are wasn't able to answer on the forum.
The first one is and I apologise if I'm mispronouncing this name, please unmute yourself and correct me if I do.
Amy Ragged, this is about discussion of matrix algebra, in particular the rank and orthogonal matrices.
The question was mainly about, how [00:17:00] does the set of properties correspond to orthogonality of matrices which is this property? What I'm going to do is I'm going to briefly discuss how these two things are linked to each other.
For that, I'm going to switch to the document camera.
Just to reposition it so that it's okay.
[00:18:00] It's maybe a bit better.
Yes, Dale? Let's look.
Indeed it is.
Very good catch.
Let me see if I can move it around.
That should be better.
Good catch, I think.
We have a matrix, a P by P matrix and it has [00:19:00] this property orthogonality, which is defined as X times X transpose, is what happens to equal X transpose X.
Which happens to be equal to an identity matrix of whatever dimension.
The question is, what does it have to do with the scalar product of two different column vectors of each of equals zero and similarly two row vectors and similarly the normal equal to 1? Which one do I want to start with? Probably the easiest way to explain this is to remember just the definition of matrix multiplication.
We have these matrices and we can draw one [00:20:00] as- let's make them three-by-three.
Now I don't know whether you've seen this little visual trick for matrix multiplication, but when you write down matrix multiplication, it often helps to put the second matrix up here.
This is transpose, so X21, [00:21:00] X31.
Then there's X12, X22, X32.
This is X and this is X transpose.
Then you have X13, X23, X33.
That's X times X transpose.
Let me see if I can maybe adjust the brightness a bit.
Is this too dark or is it okay? [00:22:00] Now we know that this has to equal 1, 1, 1, 0, 0, 0, 0, 0, 0.
Now the reason why it's helpful to write down like this is because we can say that, well, the way we can multiply it by saying, okay, we will match this value with this value, this value with this value, this value with this value to get whatever is in this cell.
Similarly we'll match this with this, this with this, this with this, to get this cell and so on.
It just makes it visually easy to see what's going on.
What do we see? Well, if this holds, then we can see that for example, XI, let me see.
[00:23:00] XIJ, so X11, X12, X13.
Then each of those could multiply by X12, X13.
This multiplied by this.
Actually let me just write it like this.
Let's say X, the diagonal element we know that that equals to the sum from J equals 1 to P of XIJ times XIJ.
From the row of the first matrix and the second [00:24:00] matrix and that's IJ.
But that's just sum J equals 1, P of XIJ^2.

Now because this has to be multiplied to identity matrix that has to equal 1.
On diagonal elements have to be one.
That means that the norm of each of the column vectors or row vectors has to be 1.
In this case, we're talking about row vectors.
This is a row vector.
This is a row vector and so on.
But if we use this property instead, we'll get column vectors.
[00:25:00] In other words, this is the squared norm of this case the ice column.
Make sense? That gets us that.
Now the scalar product of two different column vectors and row vectors equals 0.
Well, again, we can see that from here.
If we see that X transpose IJ with I does not equal J.
Again, we'll just use the definition of matrix multiplication.
I'm sorry, we need a third variable here K, XIK and then XJK.
For this element, we take these and multiply [00:26:00] it by these.
Yes jk.
For matrix, one product would be kj but here we're using x transpose and that has to equal to zero.
But notice that this is just looking at the elements of this matrix.
That's just a dot product of two different groups.
Make sense? Again, the product of the same column equals 1, the product of the different columns equals 0.
Now why is that really true orthogonality? Well, that's mapping [00:27:00] the algebra onto the geometry.
Here's a simple example.
Let's say you have your axis, x and y.
Let's say you have two points here or two vectors rather.
Again, this is the simplest possible case and if this is too small, let me know because I can draw one.
This is 0,1 vector and this is going to be 1,0 vector.
You can see they're perpendicular to each other.
In fact, if you take a dot product of them, so 0 times 1 plus 1 times 0, you'll get 0, making them orthogonal.
But for example we're perpendicular or orthogonal.
For another example, [00:28:00] let's say, we have two unit vectors that look like this.
Here this one they were to be the unit vectors, this one will be 1,1 divided by the square root of 2, so that it would have the unit length.
Then this one would be 1 minus 1 divided by square root of 2.
Then when we multiply them together, so 1 times 1 is 1, 1 times 1 times minus 1 is minus 1.
We sum them up, we get 0.
Again, these are perpendicular.
The angle between them is right angle.
This extends to any number of dimensions.
This is essentially the generalisation of the idea of perpendicular vectors.
[00:29:00] As we see here, if each of these rows, where each of these columns represents a vector in two or more dimensions, when we multiply them together we get 0.
That's how these two ideas relate.
Makes sense? Nazam? Here's the thing, I posted the solutions this morning.
Normally I will post them on Wednesday or Thursday at the very latest.
How about this? Well, actually there was a question about those solutions and I will go through one of

them.
If there's a question about the other one, I will please just attach [00:30:00] it to the next webinar thread about that.
The solutions, I've posted them, but I know it was only a few hours ago.
Next question was on the thread that I can answer.
Here is Dale Perkin and the question was about eigenvalues and eigenvectors and in particular, the example 0.
1 here.
I'm sorry, I'm still on document camera, I'm trying.
I need to switch back.
[00:31:00] It concerns this slide about spectral decomposition and in particular it concerns example 0.
1.
This one, I don't know whether I actually need a camera for.
But let's see.
Here we have a square matrix x, we're told is symmetric positive definite and has these eigenvalues Lambda 1 to Lambda p.
It has associated unit eigenvectors.
We want to show that the maximum value is attained.
[00:32:00] Maximum of this quantity that is y transpose, that's y divided by y transpose y.
So y transpose y is a scalar, so it's okay to divide by it.
The highest possible value equals to Lambda 1.
That is the biggest eigenvalue and its attained when it corresponds to the first eigenvector.
Then we can also talk about how the smallest possible value attainable is attained by this eigenvalue.
It turns out to be pretty to show and it actually has it here in the slides.
But let me walk through the reasoning here.
First because it's symmetric and positive definite, we can take the spectral decomposition of [00:33:00] it.
We can write it like this.
Now, that is as p Lambda p so we have our matrices of eigenvectors and Lambda, which is the matrix of eigenvalues.
That means we can rotate it.
Remember one property that these orthogonal vectors have or orthogonal matrices have, is that their transpose is their inverse.
By the way, that property you can actually easily see from the fact that that matrix times its transpose equals the identity matrix, which is the definition of the inverse.
Well, that means that we can essentially [00:34:00] multiply y by whatever is the inverse of the spectrum.
Let's say we define this z as p transpose y.
We can transform the problem in terms of z.
How? Well, first, we replace x by p Lambda p transpose.
Then we observe that we have this bit which is p transpose y, whole quantity transpose and we have this bit which is transpose y so that becomes z.
Now, then we can also somehow convert this y transpose y to z transpose z.
Any ideas why we can do that? [00:35:00] No worries.
That's why we're here.
Here's the thing, let's take this expression, z equals p transpose y and solve for y.
Well, one way we can do that is by pre-multiplying both sides by P transpose inverse, which is just P.
That means that we can write y equals P times z or P.

Make sense? I can switch to the document camera and write it out, but that means I have to switch back, so it's a bit of a hustle, but is there [00:36:00] anybody to whom that does not make sense? That y equals Pz.
So y equals Pz.
So let's substitute that in the denominator here.
I'm selecting.
If you'd like, I can go and do this on the paper, but it requires switching things between things.
So here this y just becomes Pz, Y transpose becomes z transpose P transpose.
What we have is z transpose P transpose Pz.
Well, remember P is an orthogonal matrix, so P transpose p is the identity matrix, so it just goes away, leaving only Z [00:37:00] transpose z.
But Lambda here is a diagonal matrix.
So we just have Z transpose to some diagonal matrix times z, but that we can write just by definition of matrix multiplication as essentially Lambda i, Zi, and then summed up, because all the other elements of Lambda are zero.
We can write it like this.
Now the denominator, we just have to get the definition of it, which is the sum of Zi squared.
Again, definition of matrix multiplication.
Makes sense? [00:38:00] So the question is then, we can make Z anything we want, right? But we want to make this ratio here as big as possible.
No w, the way we can do it is by making the Zi that gets multiplied by the biggest number here, as big as possible.
The rest we can just set to zero because they don't give us anything.
Now, we know that Lambda 1 is the biggest of the Lambdas.
Now, which means that we can maximise this quantity by [00:39:00] just setting z1 to say one and set all the others that Zis to zero.
Then the denominator becomes one, the numerator becomes Lambda i.
That make sense? The other way to write this, and this is a more formal way to write this is, well, since all the Lambda is are smaller or less than or equal to Lambda 1, we can just replace each of these with Lambda 1 then the result will be greater than or equal to anything, to whatever this might be, with us getting us our maximum.
[00:40:00] Then we can show that it's achieved here.
If we want to make this ratio as small as possible, we will want the Zi to be one that corresponds to the smallest of the Lambdas, which is Lambda p.
Then of course one last step is this.
So we know our our Zi, so for the maximisation case is going to be essentially vector with a one followed by a bunch of zeros.
Well, what happens when we translate that back to y? Now remember that y is p times z.
What happens [00:41:00] when we take that product? Well, it's going to be p multiplied by that, which will give us the first eigenvector.
If you'd like, I can go through any part of that on the document camera.
So principal component analysis, that's Week 2, so we'll get there.
Principal component analysis is about maximising the direction of variation or figuring out in what direction [00:42:00] of data variation is maximised and it turns out to come up with that form.
But also it's an exercise just to get you at least somewhat comfortable with these eigenvectors and eigenvalues and all that.
Any other questions? So the next pre-scheduled item was a question by Larissa Simpson, and this concerned the check your understanding exercise [00:43:00] number 2.

In particular, it's this one.
Now, I've posted the solution to it, so let me just pull up the relevant document here.
[00:44:00] Here's the document.
Now let me share the relevant slide here.
Where's that Window? I want to see if I need to.
That's better.
[00:45:00] This is the solution.
Now there's a number of ways to solve this.
First of all, we're just to summarise we're given some mean and some variance-covariance matrix.
Now, the thing here is that it has a particular structure, which is that the variance of X and the variance of Y is the same.
Or the variance of X_1 [00:46:00] and X_2 is the same.
It's Sigma squared and then the correlation is Rho.
That's what this signifies.
Make sense? The first thing is we want to know about the covariance between X_1 and X_2 and the sum of X_1 and X_2.
Again, there are a number of ways to do this.
The one that's I'm suggesting here is to write this as a linear transformation.
Let's see, I think might [00:47:00] be in the notes because I think we haven't had a slide about covariances so I think may have been lost.
You know what? Let me see if I can just go work through this.
The idea is that we can represent Y, which is the transformed variable.
Well, so let's call this X minus 1, also X_1 minus X_2, and X_1 plus X_2.
Let's call those two put together, stuck together.
Let's call them Y, the vector Y.
Then we can write this Y as [00:48:00] some matrix times the original X.
You can see that.
If you view X as X_1 and X_2, well, you have 1 times X_1 plus minus 1 times X_2.
That gets you the first element of Y.
Then you have their sum, which is just 1 times X_1 plus 1 times X_2.
I thought I had this in the notes somewhere, if not, then again, I need to fix that.
Generally, a variance-covariance matrix of some vector or some constant matrix times a vector is just going to be that matrix times the variance of that vector times that matrix transpose.
[00:49:00] Once we multiply this out, we will in fact get this expression.
Now, one interesting thing is if things end up cancelling out, so this covariance is always going to be 0.
Now, because this matrix is going to be symmetric, we only need to compute this for one combination only.
This is it.
Now, and again, I thought that was a notion.
In fact, maybe let me switch document camera and write this down.
This is a pretty standard [00:50:00] property of variances if you have a matrix A, it's a q by p matrix, and it's a constant.
Then you have X is a real vector of dimension p with its expected value of X equals and its variance.
[00:51:00] Then it's always true that the expected value of AX, is a A_m, and the variance or covariance matrix of AX is A, Sigma A transpose.
Sorry about that.
Apologies for the messy writing.
That also, so this is the properties here.

There's another way to do this using the bi-linearity of covariance, which I believe goes something like this.
[00:52:00] Cov, X_1, X_1 minus Cov X_1,X_2 minus Cov X_2,X_1 minus [00:53:00] Cov, X_2, X_2.
Now, so here basically we can split things up like that.
We could take X_1 and X_1, take their covariance, and X_2, X_1 and this is minus sign, so minus their covariance, and so on.
Then when we plug in the values, and there is some problem.
We end up with, excuse me, you have to write here Sigma squared times X_1, X_1, that's just 1, minus, and then you have X_1, X_2, that's going to be Rho, but again, we're factoring out the Sigma here plus covariance of X_1 and X_2, that's Rho, and then minus the covariance of X_2 and X_2.
[00:54:00] That's 1.
Again, we get cancel zero.
That's another way to use this property of covariances.
Does that make sense? By the way, the covariance of a variable with itself is just it's variance.
Then, the next part of that question is, what is the covariance between X_1 and X_2 minus Rho X_1.
Here, again, we can use either of these methods.
By just using matrix multiplication approach, we can easily get zero or using the by linearity, we can use zero.
Now last question is, [00:55:00] of course, let's say, that it's a Rho but some other B.
Then instead of covariance, we want to minimise the variance.
Again, one way you do that is, you have the variance, well, we again write this as a linear combination.
We observe that we can write a matrix capital.
Sorry about that, I meant to show the screen.
For Part B, again, you write this as a linear combination, and that's what was that.
[00:56:00] For the third part, which looks intimidating, but really we just define a to be minus b plus 1 to get minus bX_1 plus X_2.
When we multiply all that together, we end up with this expression.
It's a nice quadratic expression, which means that we can take a derivative, and observe that B equals Rho at equivalent minimum.
We know that it's a minimum because the second derivative is always positive.
The funny thing is, in principle, given the results from Part B, we don't actually need to do any of this, because, [00:57:00] well, do we need to do that? Sorry, I don't think is there.
Actually, we might need you to still do it.
We might need to at least plug-in Rho here to see that zero, and then you can't have a variance less than zero, so there you go.
What was the next item? Next question was? Yes, so I think, these are all the questions that we're asking in thread, that I didn't answer in the thread, I think.
We started about 40 min, so what I will do is, just answer questions ad [00:58:00] lib.
We'll talk about whatever thing you want me to talk about.
If we're idle for a bit, then these quizzes they don't have to be done in one sitting.
If there's a timer, then the timer does start running when you first view the quiz.
However, if there's no time limit, then you can leave and come back anytime.
Any other questions? [00:59:00] What I'm going to do is, I'm going to shift and do some consultation mode.
Which means that I'm going to put up this little bit of that.
Basically, I'm going to mute myself, and turn off the video.
If you have any questions, please raise your hand, I'm monitoring the chat.
But this way, please I can get up and stretch.

[01:00:00] Yes.
The solutions to the challenges.
That's a good question.
The solution to the challenges are, actually, there's not the same on a different tab in that screen.
I'll release those again, later in the week.
[01:01:00] Akromo, and again, by the way, if I'm mispronouncing your name, please unmute yourself and tell me no uncertain terms.
Can you come back to the quiz? Yes.
Yes, you can.
There might be an overall time limit but you can, generally in a Moodle quiz, it saves your progress.
So you can come back to the quiz anytime before it's due.
[01:02:00] Typically these started [01:03:00] 5 past [inaudible 01:03:09] and end at [01:04:00] [01:05:00] [01:06:00] [01:07:00] [01:08:00] [01:09:00] [01:10:00] [01:11:00] [01:12:00] [01:13:00] [01:14:00] [01:15:00] [01:16:00] [01:17:00] [01:18:00] [01:19:00] [01:20:00] [01:21:00] [01:22:00] [01:23:00] [01:24:00] [01:25:00] [01:26:00] [01:27:00] [01:28:00] [01:29:00] [01:30:00] [01:31:00] [01:32:00] [01:33:00] [01:34:00] [01:35:00] [01:36:00] [01:37:00] [01:38:00] 5 to the hour.
I'm going to go ahead and log out.