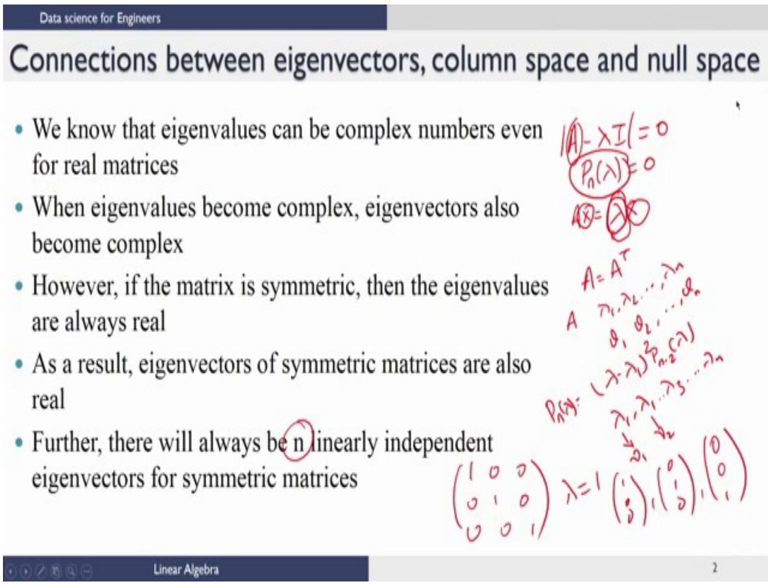
Data Science for Engineers

Prof. Raghunathan Rengasamy
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture - 18
Linear Algebra - Distance, Hyperplanes and Halfspaces, Eigenvalues, Eigenvectors

This is the last lecture in the series of lectures on Linear Algebra for data science and as I mentioned in the last class, today, I am going to talk to you about the connections between eigenvectors and the fundamental subspaces that we have described earlier. We saw in the last lecture that the eigenvalue eigenvector equation results in

(Refer Slide Time: 00:40)



this equation having to be satisfied which is A - λ I = 0. In general, we also saw that, this would turn out to be a polynomial of degree n in λ, which basically means that even if this matrix A is real, because the solutions to a polynomial equation could be either real or complex, you could have eigenvalues that are complex.

So, for a general matrix, you could have eigenvalues which are either real or complex. And notice that since we write the equation Ax = λ x, whenever this eigenvalues become complex, then the eigenvectors are also complex vectors. So, this is true in general;

however, if the matrix is symmetric and symmetric matrices are of the form $A = A^T$, then there are certain nice properties for these matrices which are very useful for us in data science. We also encounter symmetric matrices, quite a bit in data science for example, the covariance matrix turns out to be a symmetric matrix and there are several other cases where we deal with symmetric matrices.

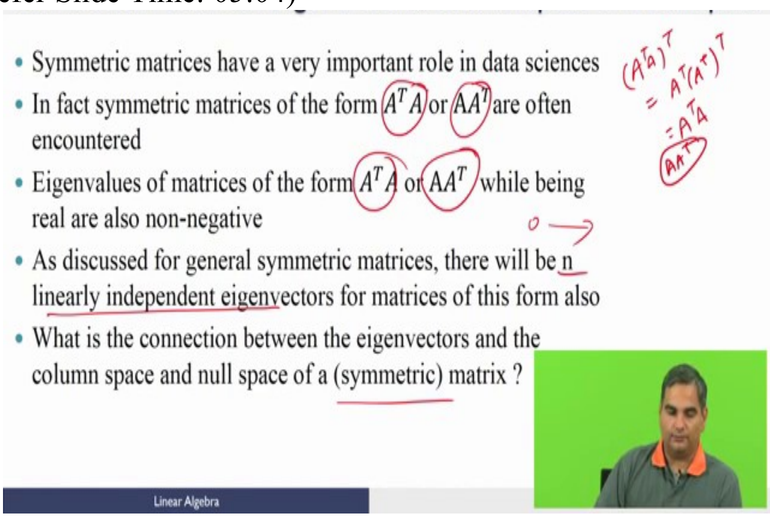
So, these properties of symmetric matrices are very useful for us when we look at algorithms in data science. Now, the first property of symmetric matrices that is very useful to us is; if the matrix is symmetric, then the eigenvalues are always real. So, irrespective of what that symmetric matrix is, this polynomial would always give real solutions for symmetric matrices. And as I mentioned before if this turns out to be real, then the eigenvectors are also real. Now, there is another aspect of an eigenvalues and eigenvectors that is important; if I have a matrix A and I have n different eigenvalues λ_1 to λn, all of them are distinct, then I will definitely have n linearly independent eigenvectors corresponding to them which could be v_1; v_2 all the way up to vn; however, if there are certain eigenvalues which are repeated. So, for example, if we take a case where eigenvalue λ_1 is repeated, then I could have some polynomial, which is like this.

So, the polynomial, original polynomial has eigenvalue; λ_1 repeated twice and then there is another n - 2 order polynomial, which will give you n - 2 other solutions. Now, in this case, when I have λ_1 repeated like this, then it could turn out that this eigenvalue either has 2 eigenvectors, which are independent or it might have just one eigenvector.

So, finding n linearly independent eigenvectors is not always guaranteed for any general matrix and we already know that eigenvectors could be complex for any general matrix; however, when we talk about symmetric matrices, we can say for sure that the eigenvalues would be real, the eigenvectors would be real, further we are always guaranteed that we will have n linearly independent eigenvectors for symmetric matrices. It does not matter how many times the eigenvalues get repeated. One classic example of a symmetric matrix, where eigenvalues are repeated many times, so take identity matrix, something like this here, this identity matrix has eigenvalue $\lambda = 1$, which is repeated thrice.

But, it would have three independent eigenvectors; 1, 0, 0; 0, 1, 0 and 0, 0, 1. So, this is a case where eigenvalues repeated is thrice, but there are three independent eigenvectors. So, this is also an important result that we should keep in mind.

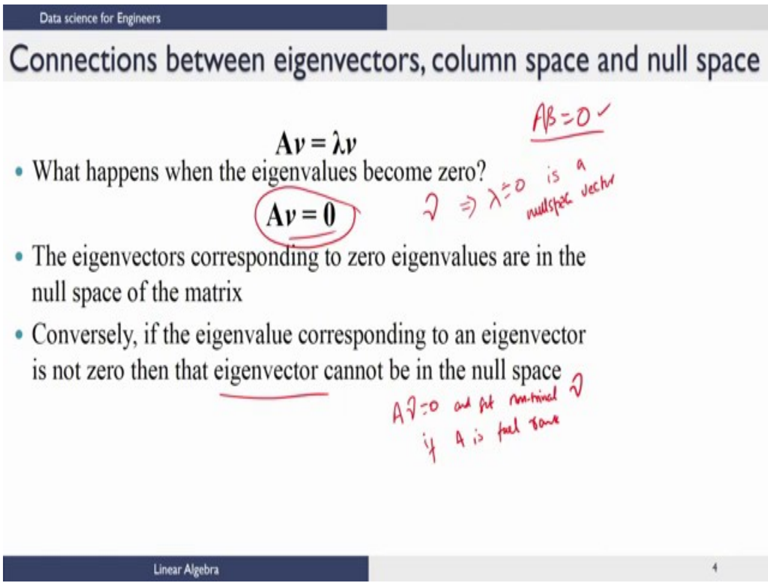
(Refer Slide Time: 05:04)

1. 

And as I mentioned in the last slide, Symmetric matrices have a very important role in data sciences. In fact, symmetric matrices of the type, $A^T A$ or AA^T are often encountered in data sense computations. And notice that both of these matrices are symmetric. So, for example, if I take $A^T A^T$, this will be A^T; A^T, which will be $A^T A$. So, the transpose of the matrix is the same. You can verify that, AAtranspose is also symmetric through the same idea. So, we know matrices of the form $A^T A$ or AA^T are both symmetric and they are often encountered; when we do computations in data science. And we know from the previous slide, I had mentioned that the eigenvalues of symmetric matrices are real, if the symmetric matrix also takes this form or this form.

We can also say that while the eigenvalues are real; they are also non-negative, that is they will be either 0 or positive, but none of the eigenvalues will be negative. So, this is another important idea that we will use; when we do data science, when we look at covariance matrices and so on. Also the fact that, this $A^T A$ and A^T are symmetric matrices; guarantees that there will be n linearly independent eigenvectors for matrices of this form also. So, what we are going to do right now is, because of the importance of symmetric matrices in data science computations, we are going to look at the connection between the eigenvectors and the column space a null space for a symmetric matrix. Some of these results translate to non-symmetric matrices also, but for symmetric matrices, all of these are results that we can use.

(Refer Slide Time: 07:04)

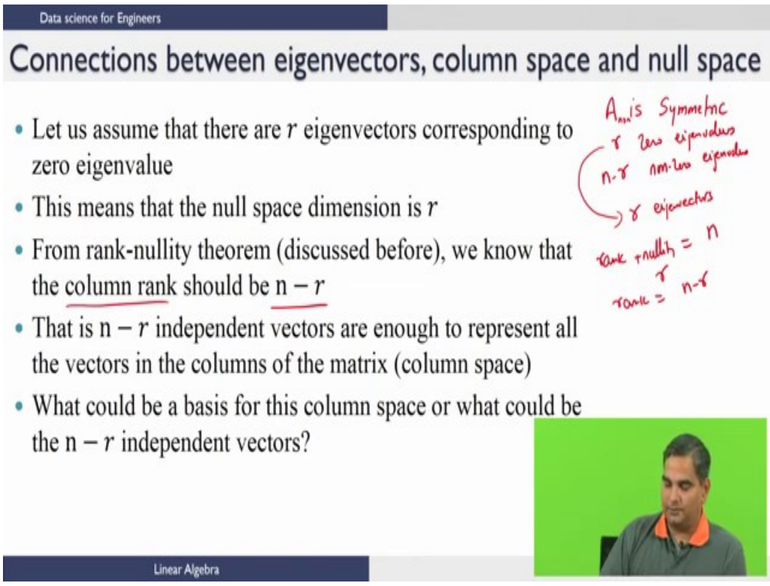


So, we go back to the eigenvalue eigenvector equation; Aν is λ ν. And this result that we are going to talk about right now, is true whether the matrix is A symmetric or not. If $A\nu = \lambda\nu$, we ask the question, what happens when λ is 0? That is one of the eigenvalues becomes 0. So, when one of the eigenvalues becomes 0, then we have this equation which is $A\nu = 0$. So, we can interpret ν as an eigenvector corresponding to eigenvalue 0.

We have also seen this equation before, when we talked about different sub-spaces for matrices; we saw that null space vectors are of the form A, β is 0 from one of our initial lectures. You notice that, this and this form are the same. So, that basically means that, ν which is an eigenvector corresponding like corresponding to eigenvalue, λ = 0, is a null space vector, because it is just of the form that we have here . So, we could say, the eigenvectors corresponding to 0 eigenvalues are in the null space of the original matrix A. Conversely, if the eigenvalue corresponding to an eigenvector is not 0, then that eigenvector cannot be in the null space of A. So, these are important results that we need to know.

So, this is how eigenvectors are connected to null space. If none of the eigen-values are zero, that basically means that the matrix A is full rank and; that means, that I can never solve A ν = 0; and get non trivial ν . So, it is not possible, if A is full rank. So, if A is full rank, I cannot solve for this and get non trivial ν. So, whenever λ is; lambdas are such that, there are there is no eigenvalue that is zero; that means, a is full rank matrix; that means, there is no eigenvector such that a ν is 0 which basically means that there are no vectors in the null space.

(Refer Slide Time: 09:40)



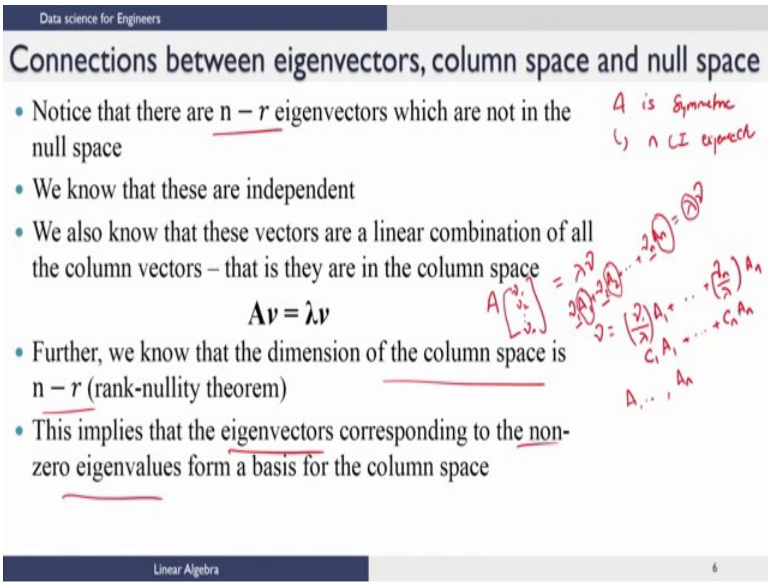
Now, let us see the connection between eigenvectors and column space. In this case, I am going to show you the result; and this result is valid for symmetric matrices. Let us assume that I have a symmetric matrix A; and the symmetric matrix A, we know will have n real eigenvalues. Let us assume that r of these eigenvalues are 0.

So, this r could be 0 also; that means, there is no eigenvalue which is zero. So, even then all of this discussion is valid. But as a general case, let us assume that r eigenvalues are 0. So, there are r zero eigenvalues. And since we are assuming this matrix is n by n, there will be n real eigenvalues of which r are 0. So, there will be n - r non-zero eigenvalues. And from the previous slide, we know that the r eigenvectors corresponding to this r 0 eigenvalues are all in the null space ok. So, since I have r 0 eigenvalues, I will have r eigenvectors corresponding to this.

So, all of these r eigenvectors are in the null space which basically means that the dimension of the null space is r; because there are r vectors in the null space; and from rank-nullity theorem, we know that rank + nullity = number of columns in this case n; since there are r eigenvectors in the null space, nullity is r. So, the rank of the matrix has to be = n - r.

So, that is what we are saying here. And further we know that column rank = row rank; and since the rank of the matrix is n - r, the column rank also has to be n - r. This basically means that there are n - r independent vectors in the columns of the matrix. So, one question that we might ask is the following; we could ask what could be a basis set for this column space? Or what could be the n - r independent vectors that we can use as the columns subspace?

(Refer Slide Time: 12:24)



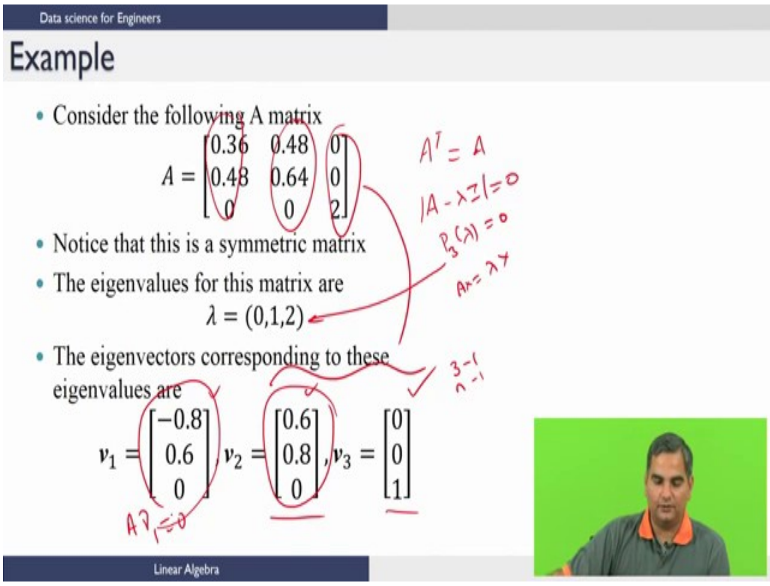
So, there are a few things that we can notice based on what we have discussed till now. First, notice that the n - r eigenvectors that we talked about in the last slide, the ones that are not eigenvector is corresponding to $\lambda = 0$; they cannot be in the null space; because λ is a number which is different from 0. So, these n - r eigenvectors cannot be in the null space of the matrix A. So, let me write again. We are discussing all of this for symmetric matrices. We know, that all of this n - eigenvectors are also independent; because we said irrespective of what the symmetric matrix is, we will always get n linearly independent eigenvectors.

So, that means, these n - r eigenvectors are also independent. We also know that each of these independent eigenvectors are going to be linear combinations of columns of A. To see this, let us look at this equation. So, let me write this out. So, I could call this as A, I am going to expand this v, into v_1, v_2 , all the way up to vn; notice that these are components of v. We are just taking one eigenvector v and then these are the n components in that eigenvector. I can write this as λ v and from the previous lecture of how to do this column multiplication and how to interpret this column multiplication, I said you could think of this as v_1 times the first column of A + v_2 times the second column of A; all the way up to vn types; nth column of A = λ v. Now in this equation, let me be very clear; these are scalars which are components in the eigenvector v; these are column vectors; this is a first column of A, second column of A, this is nth column of A, this is again a scalar λ; which is the eigenvalue corresponding to v.

So, this could be true for any of the n - r eigenvectors; which are not in the null space of this matrix A. Now, take λ to the other side. So, you will have this equation as v, is v_1 by λ A_1 and so on + vn by λ A_n. Again v_1 is a scalar λ is a scalar. So, these are all constants that we are using to multiply these columns. Now you will clearly see that, each of these eigenvectors; n - r eigenvectors are linear combinatins of the columns of A. So, there are n - r linear linearly independent eigenvectors like this and each of this are combinations of columns of A. And we also know that the dimension of the column space is n - r. In other words, if you take all of these columns, A_1 to A_n; these can be represented using just n - r linearly independent vectors.

Now, when we put all of these facts together, which is the n - r eigen-vectors are linearly independent; they are combinations of columns of A; and the number of independent columns of A can be only n - r. So, this implies that the eigenvectors corresponding to the non-zero eigenvalues for a symmetric matrix form a basis for the column space. So, this is the important result that I wanted to show you, with all of these ideas. Now again these results we will see and use as we look at some of the data science algorithms later.

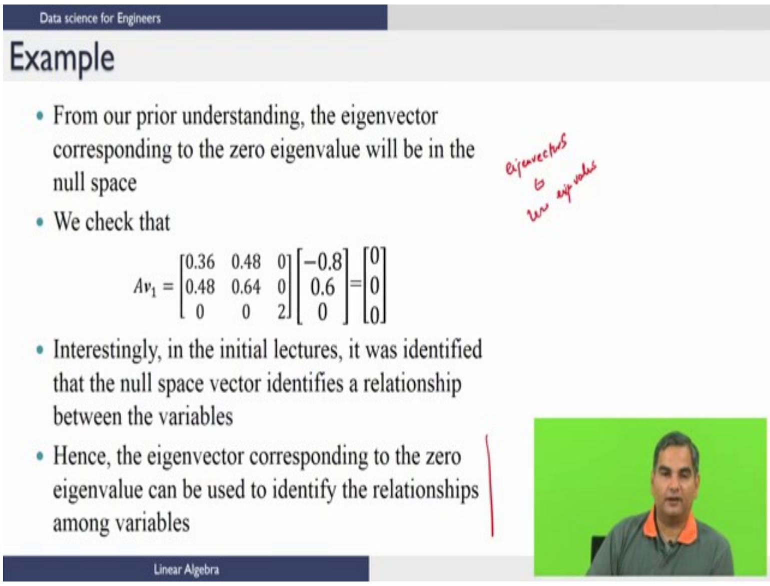
(Refer Slide Time: 16:20)



So, let us take a simple example to understand how all of these work. Let us consider a matrix which is of this form here; it is a 3 by 3 matrix. First thing that I want you to notice, that this is a symmetric matrix. So, if you do $A^T = A$. And we said symmetric matrices will always have real eigenvalues and when you do the eigenvalue computation for this, the way you do the eigenvalue computation is, you take determinant $A - \lambda I = 0$, then you are going to get a third order polynomial , you set it = 0; and then you calculate the three solutions to this polynomial and these would turn out to be the solution 0, 1, 2 and you take each of these solutions and then substitute it back in and then solve for $Ax = \lambda x$.

Then you would get the three eigenvectors corresponding to this which are given by this, this and this. I have noticed from our discussion before; since this is an eigenvector corresponding to $\lambda = 0$; so, this is going to be in the null space of this matrix A and these are the remaining 2; how do I get this 2? Which is 3 - 1, n, n by n. So, it is A_3 by 3 matrix and nullity is 1; because there is only one eigenvector corresponding to $\lambda = 0$. So, I get 2 other linearly independent vectors. And in the last slide, when we were discussing the connections, we claim that these two eigenvectors will be in the column space or in other words, what we are claiming is that these three columns can simply be written as a linear combination of these two columns; and we are also sure that when we do A times v_1, this will go to 0. So, let us verify all of this in the next slide.

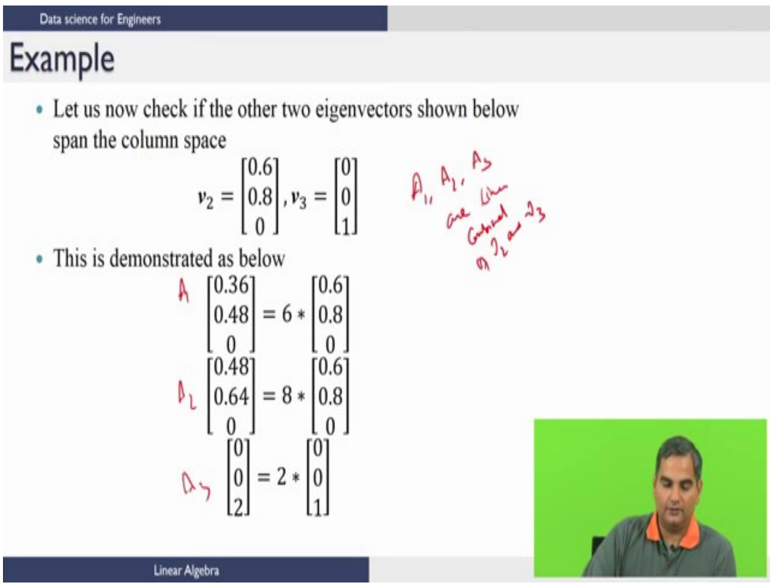
(Refer Slide Time: 18:26)



So, let us first check A times v_1. So, this is a matrix, I have a times v_1 here and you can quite easily see that when you do this computation, you will get this 0, 0, 0; which basically shows that this is the eigenvector corresponding to zero eigenvalue. Interestingly, in our initial lectures, we talked about null space and then we said the null space vector identifies a relationship between variables. Now, since this eigenvector is in the null space, the eigenvector corresponding to the eigenvector, corresponding to zero eigenvalue or eigenvectors corresponding to zero eigenvalues, identify the relationships between the variables because these eigenvectors are in the null space of the matrix.

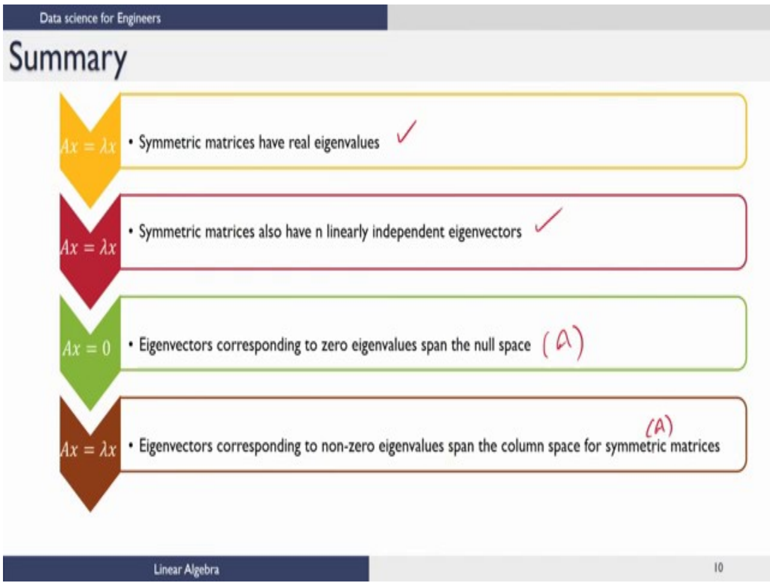
So, it is an interesting connection that we can make. So, the eigenvectors corresponding to zero eigenvalue can be used to identify relationships among variables.

(Refer Slide Time: 19:33)



Now, let us do the last thing that we discuss. Let us now check the other two eigenvectors shown below. So, this is for the other two eigenvalues, span the column space. So, what I have done here is, I have taken each one of these columns from matrix A. So, this is column 1. So, this is A_1, this is A_2 and this is A_3. Column 1 is 6 times v_2, column 2 is 8 times v_2 and column 3 is 2 times v 3. So, we can say that A_1 A_1, A_2 and A_3 are linear combinations of v_2 and v 3.

(Refer Slide Time: 20:29)



So, v_2 and v 3 form a basis for this column space of matrix A. So, to summarize, we have Ax = λ x and we largely focused on symmetric matrices in this lecture. So, we saw that, if we have symmetric matrices, they have real eigenvalues. We also saw that symmetric matrices have n linearly independent eigenvectors. We saw that the eigenvectors corresponding to zero eigenvalues span the null space of the matrix A and eigenvectors corresponding to nonzero eigenvalues span the column space of A for symmetric matrices that we described in this lecture. So, with this, we have described most of the important fundamental ideas from linear algebra that we will use quite a bit in the material that follows.

The linear algebra parts will be used in regression analysis, which you will see as part of this course. And many of these ideas are also useful in algorithms that do classification for example, we talked about half spaces and so on; and the notion of eigenvalues and eigenvectors are used pretty much in almost every data science algorithm, of particular node is one algorithm which is called the principle component analysis which we will be discussing later in this course where these ideas of connections between null space, column space and so on are used quite heavily.

So, I hope that we have given you a reasonable understanding of some of the important concepts that you need to learn to understand some of the material that we are going to teach in this course and as I mentioned before, linear algebra is a vast topic. There are several ideas; how, how do these ideas translate, which ones of these are applicable or not applicable for non-symmetric matrices and so on. And from the previous lectures, how do we develop some of those concepts more can be found in many good linear algebra books; however, our aim here has been to really call out the most important concepts that we are going to use again and again in this first course on data science for engineers, more advanced topics in linear algebra will be covered when we teach the next course on machine learning where those concepts might be more useful in advanced machine learning techniques that we will teach. So, with this we close this series of lectures on Linear Algebra and the next set of lectures would be on the use of statistics in data science.

I thank you and I hope to see you back after you go through the module on statistics which will be taught by my colleague professor Shankar Narasimhan.