Linear Algebra
Week 12

G-07

12 XII 2024

1 Eigenvalues and Eigenvectors

Eigenvalues tell us a lot about a matrix. Their power and use cases are presented in the lecture notes by developing a closed formula for Fibonacci sequence which is a quite good example. But let us start directly with the definition.

1.1 Definition and Properties

Definition 7.1.1. Given $A \in \mathbb{R}^{n\times n}$, we say $\lambda \in \mathbb{C}$ is an eigenvalue of A and $v \in \mathbb{C}^n \setminus \{0\}$ is an eigenvector of A, associated with the eigenvalue λ, when the following holds:

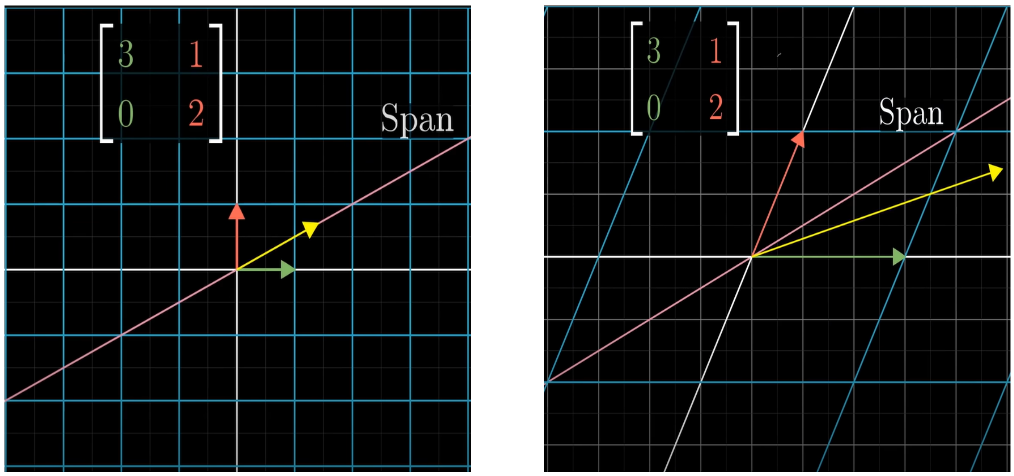
$$Av = \lambda v$$

We call them an eigenvalue-eigenvector pair. If $\lambda \in \mathbb{R}$ then we will call λ a real eigenvalue, and the associated eigenvalue-eigenvector pair a real eigenvalue-eigenvector pair.

What does $Av = \lambda v$ actually mean? Geometrically this means, when applied the linear transformation corresponding to the matrix A is equivalent to scaling the vector v. 3Blue1Brown visualizes this in the video on eigenvalues.[1]

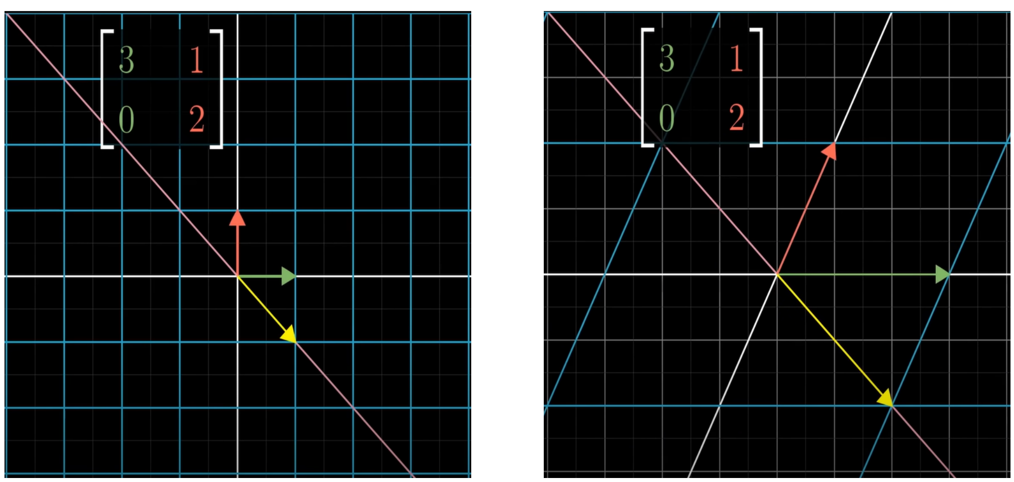
[1] I will put some snapshots here for you to have the picture in mind, but watching the whole video is strongly recommended. *By Grant Sanderson, from https://youtu.be/PFDu9oVAE-g?si=Qh0dtPO5YTaUzZOi*

If a vector is not an eigenvector, it gets knocked off its span when the linear transformation is applied. In other words, after multiplication with the matrix, the vector is not on the same line as before. Below, red and green vectors are the standard unit vectors to show the general effect of the transformation.



before linear transformation   after linear transformation

v is NOT an eigenvector



before linear transformation   after linear transformation

v IS an eigenvector

If our vector is an eigenvector of the matrix associated with the linear transformation, then the effect of multiplying the vector with that matrix is the same as multiplying the vector with a scalar which we call an eigenvalue. In other words eigenvectors stay on the same line as before, when multiplied with the matrix. Many properties of the eigenvalues and eigenvectors are more intuitive if you visualize them this way, so keep this in mind for later.

If we inspect the equality $Av = \lambda v$ more, we can put it in some fancy linear algebra language. We have:

$$
\begin{aligned}
Av &= \lambda v \\
Av - \lambda v &= 0 \\
(A - \lambda I)v &= 0
\end{aligned}
$$

This means if λ and v are an eigenvalue-eigenvector pair, then the vector v must be a nonzero[2] element of the Nullspace of the matrix $A - \lambda I$. We can then say this matrix is not invertible because it has a non trivial nullspace. This implies $det(A - \lambda I) = 0$. We can and will exploit this fact when we try to calculate eigenvalues and eigenvectors of a matrix.

How to calculate eigenvalues & eigenvectors:

1. *Simplify $det(A - \lambda I) = 0$ and leave λ as a variable. This expression becomes a polynomial in λ. Solve it. You have the eigenvalue(s).*

2. *Now you have the eigenvalues. Solve $Av = \lambda_i v$ for all λ_i you found in the previous step.*[3] *Be careful $v \neq 0$ must hold per definition. Now you also have eigenvectors.*

Note that there is only a finite set of eigenvalues but there are infinitely many eigenvectors! If you scale each equation in the LSE from the 2. step, then it still holds. A natural question is, is there always at least an eigenvalue for all matrices? The answer is yes! Since we have a polynomial in the 2. step the

[2]nonzero by definition, see Def. 7.1.1.

[3]It is simply a linear system of equations with n unknowns which are the elements of the vector v $(v_1, v_2, \dots v_n)$ and n equations, because our matrix must be square by definition.

fundamental theorem of algebra implies that there will always be a solution, maybe complex valued.

The following propositions and theorem formalize what we discussed above.

Proposition 7.1.2. Let $A \in \mathbb{R}^{n \times n}$. $\lambda \in \mathbb{R}$ is a (real) eigenvalue of A if and only if $det(A - \lambda I) = 0$. A vector v is an eigenvector associated with the eigenvalue λ if (and only if) it is a non-zero element of $\mathbf{N}(A - \lambda I)$.

Proposition 7.1.3.$det(A - \lambda I)$ is a polynomial, in and λ, of degree n. The coefficient of λ^n term is $(-1)^n$.

(By the formula for the determinant (Def 6.0.6))

Theorem 7.1.4. Every matrix $A \in \mathbb{R}^{n \times n}$ has an eigenvalue (maybe complex-valued).

(Directly implied by the Fundamental Theorem of Algebra)

Note that all of these definitions hold in $\mathbb{C}$ as well. You just replace $\mathbb{R}$ with $\mathbb{C}$ and do the necessary adjustments.

Now we go through some properties of the eigenvalues.

Proposition 7.1.6. If λ and v are an eigenvalue-eigenvector pair of a matrix A, then, for $k \geq 1$, λ^k and v are an eigenvalue-eigenvector pair of the matrix A^k.

Proof: There is a short induction proof in the lecture notes where the induction step is $A^k v = A(A^{k-1}v) = A(\lambda^{k-1}v) = \lambda^k v$ and the base case for $k = 1$ is trivial.
Geometrically: Intuitively you can think of multiplying with A as scaling v. If you do this two times in a row, the first time you are scaling v by λ second time you are scaling λv by λ and therefore have $\lambda^2 v$.

Proposition 7.1.7. Let A be an invertible matrix. If λ and v are an eigenvalue-eigenvector pair of the matrix A, then, $\frac{1}{\lambda}$ and v are an eigenvalue-eigenvector pair of the matrix A^{-1}.

Proof [4]*:* A is invertible and hence, in the statement $\lambda \neq 0$. Since $Av = \lambda v$ we have $A^{-1}(\lambda v) = v$ and so $\lambda A^{-1} v = v$, which (since $\lambda \neq 0$) is equivalent to $A^{-1} v = \frac{1}{\lambda} v$.
Geometrically: You can say that if the multiplication with the matrix stretches the eigenvector v by a factor of λ then the inverse transformation would shrink it by the factor of λ or stretch it with a factor of $\frac{1}{\lambda}$.

Another important property is that distinct eigenvalues have linearly independent eigenvectors.

Proposition 7.1.8. Let $A \in \mathbb{R}^{n \times n}$ and let $v_1, \ldots, v_k \in R^n$. If $\lambda_1, \ldots, \lambda_k$ are all distinct, the eigenvectors $v_1, \ldots, v_k$ are linearly independent.

You can find the beautiful proof of this proposition in the lecture notes. As a thought experiment in an example with only two eigenvectors we can argue as follows: Two nonzero vectors are linearly dependent when they are collinear. But if they are on the same line then they are both going to be scaled by a factor of λ and therefore the associated eigenvalue will be the same for both vectors. So the two dependent vectors have the same eigenvalue which is the contraposition of our proposition. *Keep in mind that this is nowhere near a formal proof.* As a consequence we have the following theorem:

Theorem 7.1.9. Let $A \in \mathbb{R}^{n \times n}$ with n distinct real eigenvalues (meaning that the n zeroes of $det(A - \lambda I)$, as described in Corollary 7.0.4 are all distinct) then there is a basis of $\mathbb{R}^n$, $v_1, \ldots, v_n$ made up of eigenvectors of A.

This special basis is sometimes also called the *eigenbasis*. Having an eigenbasis is an important property. Because if a matrix has an eigenbasis then

[4]**Note that:** We do not cancel v in the equation $A^{-1}(\lambda v) = v$. We can not do this because working with vectors and matrices is different than working with numbers. If we somehow did the cancellation which is not defined then we would have something like $A^{-1}\lambda = \mathbf{1}$ which is not possible since the dimensions of A and a vector of 1's do not match

you can do a change of basis to the eigenbasis and thereby diagonalize the matrix. We will see both concepts below but first go on with properties of eigenvalues.

We define ***the Characteristic Polynomial of a matrix*** that will help us understand some properties of the eigenvalues and eigenvectors. We have by Corollary 7.0.4.

$$(-1)^n det(A - zI) = det(zI - A) = (z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_n) \quad (1)$$

The two polynomials $det(A - zI)$ and $det(zI - A)$ actually have the same roots and both can be used to find the eigenvalues of a matrix. But most of the time - and more importantly in this course - the characteristic polynomial is defined as $det(zI - A)$, because this is a monic polynomial[5] whereas the polynomial $det(A - zI)$ is only monic when n -the dimensions of A- is even. If n is odd, then you have -1 as the leading coefficient in $det(A - zI)$.

Proposition 7.1.10. Given $A \in \mathbb{R}^{n\times n}$ the eigenvalues of A are the same as the ones of $A^\top$.

Proof [6]*:* This follows from (1), and the fact that, for $det(A - zI) = det((A - zI)^\top) = det(A^\top - zI)$.

There is also a connection between the trace and the determinant of a matrix and its eigenvalues. The details about the trace are not included here but here is the proposition that links trace, determinant and eigenvalues.

Proposition 7.1.12. Let $A \in \mathbb{R}^{n\times n}$ and $\lambda_1, \ldots, \lambda_n$ its n eigenvalues as they show up in (1) (meaning that a value λ may be repeated, the number of times it shows up is the algebraic multiplicity of the eigenvalue) then

$$Tr(A) = \sum_{i=1}^{n} \lambda_i$$

$$det(A) = \prod_{i=1}^{n} \lambda_i$$

[5]monic means the leading coefficient of the polynomial is equal to 1

[6]Be careful eigenvalues are same for A and $A^\top$ but the eigenvectors are in general not!

You can find the proof in the lecture notes. This is a useful tool to double-check your calculations quickly. Just be careful to include each eigenvalue according to their algebraic multiplicity!

Remark 7.1.15. A few important words of caution:

1. Even though the eigenvalues of A and $A^\top$ are the same, the eigenvectors are not!

2. The eigenvalues of $A+B$ are not easily computed from the eigenvalues of A and the ones of B, in particular they are not their sum!

3. The eigenvalues of AB or BA are not easily computed from the eigenvalues of A and the ones of B, in particular they are not their product!

4. Gaussian Elimination doesn't preserve eigenvalues and eigenvectors. The eigenvalues are not the diagonal elements of the U matrix in the $PA = LU$ factorization.

Proposition 7.1.17: Let $Q \in \mathbb{R}^{n\times n}$ be an orthogonal matrix. If $\lambda \in \mathbb{C}$ is an eigenvalue of Q then $|\lambda| = 1$.

Proof: The proof uses the norm preserving property of orthogonal matrices. *(Try it!)* However don't think the only possible choices are -1 and 1. In $\mathbb{C}$ these are not the only numbers with norm 1. For a real orthogonal matrix with complex valued eigenvalues you can see ***Example 7.1.16.***

1.2 Repeated Eigenvalues

The existence of a basis out of the eigenvectors of a matrix is a useful property of a matrix as I mentioned above, and will mention below. We will firther investigate this, but first, a definition:

Definition 7.1.20. If, given a matrix $A \in \mathbb{R}^{n\times n}$, we can build a basis of $\mathbb{R}^n$ with eigenvectors of A we say that A has a complete set of real eigenvectors.

That being said, We showed in ***Theorem 7.1.9.*** that if a matrix has n distinct eigenvalues then it has a complete set of real eigenvectors. How about repeated eigenvalues? This might be a problem if we want a complete set of real eigenvectors but not necessarily. Let's see the two examples from the lecture notes:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{aligned} det(A - \lambda I) &= \lambda^2 \\ \lambda_1 = 0, \lambda_2 &= 0 \\ dim(\mathbf{N}(A - 0I)) &= 1 \end{aligned} \qquad \begin{aligned} det(B - \lambda I) &= \lambda^2 \\ \lambda_1 = 0, \lambda_2 &= 0 \\ dim(\mathbf{N}(B - 0I)) &= 2 \end{aligned}$$

We calculate the eigenvalues of both matrices and find out that both have only one eigenvalue, 0, with algebraic multiplicity 2. This does not help us determine whether or not the matrices have a complete set of eigenvectors. Therefore we look at the dimension of the nullspace of the matrices $A - \lambda I$ and $B - \lambda I$, $\lambda = 0$ in both cases since it is the only eigenvalue. The reason for looking at dimension of this nullspace is because we defined an eigenvector to be a nonzero element of $N(A - \lambda I)$ for an eigenvalue of the matrix A. (***Proposition 7.1.2.***) With this knowledge we can conclude that A does not have a complete set of real eigenvalues but B does.[7]
The lecture notes perfectly summarize the situation:

"In general when there is an eigenvalue λ with algebraic multiplicity larger than 1, it can be that $N(A - \lambda I)$ is of large enough dimension to find enough linearly independent eigenvectors."

Definition 7.1.22. Given a matrix $A \in \mathbb{R}^{n \times n}$ and an eigenvalue λ of A we call the dimension of $N(A - \lambda I)$ the *geometric multiplicity* of λ.

Further Fact 42. A matrix has a complete set of eigenvectors when the geometric multiplicities are the same as the algebraic multiplicities of all eigenvalues.

[7] Indeed any vector is an eigenvector of B, so any two linearly independent vectors are a complete set of real eigenvectors for B.

Examples of matrices with a complete set of real eigenvectors with repeated eigenvalues are projection matrices. Projection matrices onto real subspaces have two eigenvalues, 0 and 1, and a complete set of real eigenvectors. Proof in the lecture notes. *(Try to prove this! Think about which vectors stay on the same line as they were before, after projection is applied.)*

2 Change of Basis

We use coordinates to describe where a point is and then pack those coordinates in a vector. If you have $v = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$, then vector v tells you to go to a point using 4 times the first basis vector plus 6 times the second basis vector. We implicitly assume these basis vectors are the standard unit vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. What if we used another basis? Take this basis: $\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix}\}$ What would the new coordinates be, in order to reach to the same point as before? Let me spoil this one, if you use 2 times the first basis vector and 1 times the second basis vector of this new basis then you would end up at the same point as before. So the vector $v = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ written in our new basis is $v' = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. And in a nutshell this is what we mean by change of basis.[8]

What did we essentially do above? How can you calculate the magical new coordinates if this example were much more complicated? Let's formalize this example to some extend. First step is to pack the new basis vectors in a matrix

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix}$$

Now we want to determine how much and which of these new basis vectors we need, in order to generate our vector v. Well this exactly corresponds to the linear system of equations $v = Ax$ or

$$\begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

[8]Nothing in this section is taken from the lecture notes. I am going to try to provide you with some intuition, but I can explain with words only so far. I again strongly recommend the corresponding 3Blue1Brown video.

Since we collected the new basis vectors in the matrix, we know the columns are linearly independent. Hence we also know that this matrix is invertible. This makes it easier to solve the LSE:

$$\left(\begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix}\right)^{-1} \begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

which gives us the new coordinates in this new basis as $v' = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, just as we calculated above. This is how you calculate the new coordinates. But what if you are given the basis and the coordinates in this basis, how do you get back to the standard basis? It is simple, you just multiply the matrix, that has the basis vectors on its columns, with the given vector. In our example if you were just given the new basis $\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix}\}$ and the vector $v' = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ written in this basis, you calculate

$$v = Av' = \begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

With the help of this, we come to the following conclusion:

Let $B = \{v_1, \ldots, v_n\}$ be a basis of $\mathbb{R}^n$ and let $A \in \mathbb{R}^{n \times n}$ be the matrix that has these basis vectors as its columns. If v' is written in basis B, then Av' gives the corresponding coordinates in the standard basis. If v is written in the standard basis, then $A^{-1}v$ gives the corresponding coordinates in the new basis.

Nothing in this section is taken from the lecture notes. In the lecture notes, you try to write a linear transformation in a new basis. Try to understand the explanations there. This will come in handy when you do Fourier Transform in the future.

3 Diagonalization

This section can be summarized in one sentence: If you write the linear transformation corresponding to the matrix A, in A's eigenbasis - basis consisting of its eigenvalues - then this transformation/matrix is simply diagonal.

If we write a vector $x \in \mathbb{R}^n$ as $x = \sum_{i=1}^{n} \alpha_i v_i$ then $Ax = \sum_{i=1}^{n} \lambda_i \alpha_i v_i$ where λ_i is the eigenvalue associated with the eigenvector v_i. Here α_i is the i^{th} coordinate of x written in the eigenbasis. Using this idea we can develop the following theorem:

Theorem 7.2.1. Let $A \in \mathbb{R}^{n\times n}$ be a matrix with a complete set of real eigenvectors (in the sense of Definition 7.1.20) and let $v_1, \ldots, v_n \in \mathbb{R}^n$ be a basis formed with eigenvectors of A and let $\lambda_1, \ldots, \lambda_n$ be the associated eigenvalues (λ_i associated to v_i). Let V be the matrix whose columns are the eigenvectors v_i,

$$V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

Then,

$$A = V \Lambda V^{-1},$$

where Λ is a diagonal matrix with $\Lambda_{ii} = \lambda_i$ (and $\Lambda_{ij} = 0$ for all $i \neq j$).

The proof is in the lecture notes. Now, in light of this theorem we define diagonalizable matrices:

Definition 7.2.2 (Diagonalizable Matrix). A matrix $A \in \mathbb{R}n \times n$ is called a diagonalizable matrix if there exists an invertible matrix V such that $V^{-1}AV = \Lambda$, where Λ is a diagonal matrix.

Here we can see that A and λ have the same eigenvalues. This is because they are similar.

Definition 7.2.3 (Similar Matrices) We say that $A \in \mathbb{R}n \times n$ and $B \in \mathbb{R}^{n\times n}$ are similar matrices if there exists an invertible matrix S such that $B = S^{-1}AS$.

Proposition 7.2.4. Similar matrices have the same eigenvalues.

4 Hints

1. Solved in class. If you want to calculate the eigenvalues, try guessing first. In most cases the question is designed in such a way that you can guess the eigenvalue-eigenvector pair.

2. For an eigenvector v of the matrix AB, consider $w = Bv$. Each subtask builds up for the next one.

3. Bonus. No hints!

4. **a)** Use $B = S^{-1}AS$ and play around with the determinant $det(B - zI)$ and $det(A - zI)$. **b)** Diagonalize both A and B. Argue why you can do that. Then multiply the diagonalized versions. **d)** Use previous subtask at some point

5. **b)** You can go to Assignment 8 Exercise 3 to get help about how to build a reflection matrix.

6. **c)** For v and λ as eigenvector eigenvalue pair you have $Av = \lambda v$. Pack v's in the columns of a matrix and λ's on the diagonal of another matrix. Write the equation again and solve for A.

7. Cute exercise, no hints.

mkilic

♠