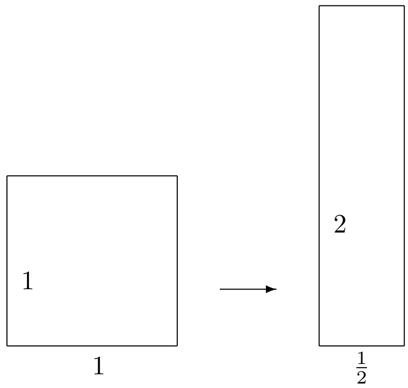
Eigenvalues and eigenvectors

Some motivation: We just saw in the last set of notes that the determinant of a 2×2 matrix tells us the effect the associated linear map has on area. In other words, if $\det([T]) = 2$ then $T : \mathbb{R}^2 \to \mathbb{R}^2$ will scale the areas of squares by a factor of 2. It's not to hard to show that T scales the areas of all shapes by the same factor. (Hint: cut whatever shape you're interested in into a bunch of little tiny squares. You won't be able to do this exactly, but what you have left over has a negligible area.) However, it's easy to find a linear map which preserves area but distorts lengths by a lot. For instance, consider the linear

$$T : \mathbb{R}^2 \to \mathbb{R}^2, \qquad [T] = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix}.$$

We draw a picture of what T does to the unit square below.



This map preserves area, but it changes lengths by a lot. It shrinks length in some directions by a factor of $1/2$ and it stretches lengths in other directions by a factor of 2. We can make this picture much worse by choosing , for instance, a horizontal scale factor of $1/100$ and a vertical scale factor of 100. This example tells us we need at least two numbers to keep track of how a linear map $T : \mathbb{R}^2 \to \mathbb{R}^2$ deforms lengths. We'll see in a bit that, at least in some special cases, we only need two numbers, and that these numbers are (essetially) the eigenvalues.

Definitions: If you know a little about the German language, you might be able to guess what an eigenvector is. The German word *eigen* means *own*, and an eigenvector of a linear transformation keeps its own direction. It can get rescaled, but the direction remains the same.

Definition 1 *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Then a nonzero vector $v \in \mathbb{R}^n$ is an eigenvector with eigenvalue λ if $T(v) = \lambda v$. Notice that, even though v is not allowed to be zero, it's possible that $\lambda = 0$.*

Exercise: Why is it necessary to have $v \neq 0$ in the definition of an eigenvector v?

This definition is a little awkward for doing computations, so the first thing we'll do is reformulate it a little. Let v be an eigenvector of T with eigenvalue λ. Then

$$[T][v] = \lambda[v] = \lambda[I][v] \Leftrightarrow ([T] - \lambda[I])[v] = 0.$$

Now, $v \neq 0$, so the linear transformation $T - \lambda I$ sends a nonzero vector to 0, which means it can't be one-to-one. This means $T - \lambda I$ isn't invertible, and so

$$\det([T] - \lambda[I]) = 0.$$

This last equation is an n-th degree polynomial equation for the unknown λ. We know that any n-th degree polynomial has exactly n roots in the complex numbers $\mathbb{C}$ (so long as we remember to count repeated roots), which means we've just proved the following

Theorem 1 *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear. Then a complex number $\lambda \in \mathbb{C}$ is an eigenvalue of T if and only if*

$$\det([T] - \lambda[I]) = 0.$$

Moreover, every $n \times n$ matrix has precisely n complex numbers $\lambda_1, \ldots, \lambda_n$ (counted with multiplicity) which are eigenvalues.

This theorem tells us how to compute eigenvalues of a square matrix: we write down the polynomial $\det([T] - \lambda[I])$ and find its roots. In practice this can be a little sticky, for instance, if we want to find the eigenvalues of a 5×5 matrix. However, for the case of 2×2 matrices, which is most of what we'll discuss in this class, the eigenvalues are the roots of a second order polynomial, which we can always find using the quadratic formula. So, for the time being at least, let's say we can find eigenvalues and continue, to see how to find the eigenvectors.

Let $[T]$ be an $n \times n$ matrix, and let λ be an eigenvalue of $[T]$. We want to find the associated eigenvector(s), that is the nonzero vectors v such that $T(v) = \lambda v$. We write this equation as a matrix equation

$$[T][v] = \lambda[v]$$

and try to solve it using our favorite method (like row reduction).

Exercise: Show that if v is an eigenvector of the matrix A with eigenvalue λ, then $2v$ is also an eigenvector of A, with the same eigenvalue λ. Is there anything special about the scale factor of 2?

Exercise: Show that the linear system $[T][v] = \lambda[v]$ for finding an eigenvector will always have many many solutions. Usually, it this system will have one free variable, so it might be convenient to set one of the components of v to 1. However, it is possible that this linear system has more than one free variables.

Ok, that's the the basics of eigenvalues and eigenvectors. You now know what they are and how to find them. For the rest of these notes we'll discuss some properties of eigenvectors and eigenvalues, and do some examples, but actually right now it might be a good idea for you to put the notes down and work through some examples. Just write down some 2×2 matrices and find their eigenvalues and eigenvectors. The first several examples will be slow and difficult, but you'll get better and faster with some practice.

Some properties of eigenvalues and eigenvectors:

Here we list some properties of the eigenvalues and eigenvectors.

Recall that v is an eigenvector of A with eigenvalue λ if $Av = \lambda v$. If λ is a real number as well, this means $A(v)$ is colinear with v, *i.e.* either $A(v)$ points in the same direction or the opposite direction as v. In other words, if λ is a real eigenvalue of A then, considered as a linear map, A preserves the direction of the associated eigenvector v.

Exercise: Recall that we constructed the 2×2 rotation matrices

$$[R_\theta] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Show that $[R_\theta]$ has a real eigenvalue if and only if the angle θ is an integer multiple of π (when measured in radians).

Exercise: We also constructed reflection matrices. Show that 1 is an eigenvalue of any reflection matrix.

Exercise: Let A be the 3×3 matrix associated to a rotation of 3-dimensional space. Show that 1 is an eigenvalue of A, and describe the relation between this associated eigenvector and the rotation.

Exercise: Show that 0 is an eigenvalue of an $n \times n$ matrix A if and only if A is not invertible. (This is completely general.)

Now we consider an $n \times n$ matrix A with real entries A_{ij} in the ith row, jth column. We have that λ is an eigenvalue of A precisely when

$$\det(A - \lambda I) = 0.$$

This is an nth degree polynomial, and the coefficients of this polynomial are sums of products of the entries of A. This means λ is a root of a polynomial with real coefficient. Now, it can happen that λ is not a real number, but it is a complex number, but these complex roots occur in conujugate pairs. We have the following

Proposition 2 *Let A be an $n \times n$ matrix with real entries. Then a non-real complex number $\lambda = a + ib$ is an eigenvalue of A if and only if its complex conjugate $\bar{\lambda} = a - ib$ is also an eigenvalue. In fact, in this case the eigenvectors are also complex conjugates. That is, if v is an eigenvector associated to the eigenvalue λ then $\bar{v}$ is an eigenvector associated to the eigenvalue $\bar{\lambda}$.*

The last sentence of the proposition follows immediately from taking the complex conjugate of the equation $Av = \lambda v$ to get $A\bar{v} = \bar{\lambda}\bar{v}$.

Exercise: Let A be an $n \times n$ matrix with real entries, and let $\lambda = a + ib$ be a non-real eigenvalue. Show that the components of the associated eigenvector v are also non-real.

Some times we can find n independent eigenvectors $v_1, \ldots, v_n$ for an $n \times n$ matrix A. This means we can find n different vectors $v_1, \ldots, v_n$ such that $A(v_j) = \lambda_j v_j$, and that we can't write v_j as the weighted sum of the other v_i's. In this case, we say that A is **diagonalizable**, for the following reason. We can write any vector w as a sum $w = c_1 v_1 + c_2 v_2 + \cdots c_n v_n$, and then

$$A(w) = A(c_1 v_1 + \cdots c_n v_n) = c_1 A(v_1) + \cdots c_n A(v_n) = c_1 \lambda_1 v_1 + \cdots c_n \lambda_n v_n. \tag{1}$$

In other words, if we change coordinates and write all our vectors as sums of $v_1, \ldots, v_n$ as above, then A has the very nice form

$$[A] = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix},$$

and (in the right coordinates) is a diagonal matrix. We see immediately from equation (1) that, at least if A is diagonalizable, that the eigenvalues $\lambda_1, \ldots, \lambda_n$ encode the stretch factors we were looking for at the beginning of these notes.

We need to know the facts that the determinant and the trace of a matrix do not depend on the basis; that is, if you change coordinates as we just did the determinant and the trace remain the same.

Exercise: Show that, for a diagonalizable, $n \times n$ matrix, the determinant is the product of the eigenvalues and the trace is the sum of the eigenvalues.

Exercise: Not all matrices are diagonalizable. In fact, show that $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not diagonalizable.

We saw above that if A is diagonalizable then $\det(A)$ is the product of all the eigenvalues and $\mathrm{tr}(A)$ is their sum. In fact, this is true for any $n \times n$ matrix, as you'll see in a second year linear algebra course when you discuss the Jordan cannonical form of a matrix.

Proposition 3 *For any $n \times n$ matrix A, it holds that $\det(A)$ is the product of the eigenvalues of A, and $\mathrm{tr}(A)$ is their sum.*

Exercise: Let A be a 2×2 matrix with complex eigenvalues $\lambda_\pm = a \pm ib$. Show that $\mathrm{tr}(A) = 2a$ and $\det(A) = a^2 + b^2$. In particular, $\det(A) \geq 0$.

Exercise: Let A be a 2×2 matrix with real eigenvalues λ_1 and λ_2. Show that $\det(A) > 0$ if and only if λ_1 and λ_2 have the same sign. Then show that λ_1 and λ_2 are both positive if and only if both $\det(A) > 0$ and $\mathrm{tr}(A) > 0$.

Finally, we mention **symmetric** matrices, that is matrices such that $A_{ij} = A_{ji}$ where A_{ij} is the entry of A in the ith row, jth column. These are particularly nice, as we see from the following theorem you'll prove in the second year linear algebra class.

Theorem 4 *A symmetric $n \times n$ matrix is diagonalizable and has n real eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \in \mathbb{R}$.*

Examples: We'll compute the eigenvalues and eigenvectors of some 2×2 matrices here, just so we have some examples written down.

First let

$$A = \begin{bmatrix} 4 & -2 \\ 3 & -3 \end{bmatrix}.$$

We want to find the eigenvalues of A, so we set

$$\begin{aligned} 0 &= \det(A - \lambda I) = \det\left(\begin{bmatrix} 4-\lambda & -2 \\ 3 & -3-\lambda \end{bmatrix}\right) \\ &= \lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2), \end{aligned}$$

and we see that the eigenvalues of A are $\lambda_1 = -2$ and $\lambda_2 = 3$.

Now we find the eigenvector associated to the eigenvalue $\lambda_1 = -2$. We want to solve the linear equation

$$Av = -2v \Leftrightarrow \begin{bmatrix} 4 & -2 \\ 3 & -3 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -2v_1 \\ -2v_2 \end{bmatrix}.$$

Of course, you can solve this using row reduction, but I find that for a small system like this, it's easier to just write out the equations. We have

$$4v_1 - 2v_2 = -2v_1, \qquad 3v_1 - 3v_2 = -2v_2,$$

and both these equations reduce to $v_2 = 3v_1$. (You might want to think about why you'll always reduce from two equations to one when you're finding the eigenvectors of a 2×2 matrix.) So, up to a scale factor, the eigenvector of A associated to $\lambda_1 = -2$ is

$$v = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Finally we find the eigenvector associated to $\lambda_2 = 3$. This time the linear equation is

$$Aw = 3w \Leftrightarrow \begin{bmatrix} 4 & -2 \\ 3 & -3 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 3w_1 \\ 3w_2 \end{bmatrix},$$

which we rewrite as

$$4w_1 - 2w_2 = 3w_1, \qquad 3w_1 - 3w_2 = 3w_2.$$

This reduces to $w_1 = 2w_2$, and so the eigenvector is (again, up to scale)

$$w = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

For our next example, we take the matrix

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Again, we find eigenvalues of A by setting

$$\begin{aligned} 0 &= \det(A - \lambda I) = \det\left(\begin{bmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{bmatrix}\right) \\ &= \lambda^2 - 2\lambda + 2. \end{aligned}$$

Using the quadratic formula we see that the eigenvalues are $\lambda_+ = 1 + i$ and $\lambda_- = 1 - i$. Notice that, just as we said earlier, the eigenvalues occur in conjugate pairs.

We set up the equation for the eigenvector associated to $\lambda_+ = 1 + i$ as before, and get

$$Av = (1+i)v \Leftrightarrow \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} (1+i)v_1 \\ (1+i)v_2 \end{bmatrix},$$

which reduces to

$$v_1 = iv_2 \Leftrightarrow -v_2 = iv_1.$$

(You might want to recall here that $\frac{1}{i} = -i$.) So we see that the eigenvector associated to $\lambda_+ = 1 + i$ is

$$v = \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

There's a short cut to finding the other eignevector w: since

$$Aw = \lambda_- w = \bar{\lambda}_+ w$$

and $\bar{A} = A$ we must have

$$w = \bar{v} = \begin{bmatrix} -i \\ 1 \end{bmatrix}.$$

We can also do this computation directly:

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} (1-i)w_1 \\ (1-i)w_2 \end{bmatrix} \Leftrightarrow w_2 = iw_1 \Leftrightarrow w_1 = -iw_2$$

and we recover

$$w = \begin{bmatrix} -i \\ 1 \end{bmatrix}.$$

An application to dynamical systems: Often one can represent a discrete dynamical system by an $n \times n$ matrix A. (Here we discretize the time variable t.) We can think of A as a recipe for how the system evolves from one time step to the next. That is, if v describes the state of the system at time t_n then $A(v)$ describes the state of the system at time t_{n+1}. Now suppose v is an eigenvector of A with eigenvalue λ, and we begin our dynamical system at the position v at time $t_0 = 0$. In the next time step we evolve to position $A(v) = \lambda v$, and in the time step after that we evolve to

$$A(A(v)) = A(\lambda v) = \lambda A(v) = \lambda^2 v,$$

and so on. After N time steps, we arrive at position $A^N(v) = \lambda^N v$. Now, if $|\lambda| < 1$ we have $|\lambda|^N \to 0$ as $N \to \infty$, so the long-time behavior of this system, at least if we start at the eigenvector v, is to collapse down to zero. On the other hand, if $|\lambda| > 1$ then $|\lambda|^N \to \infty$ as $N \to \infty$, and so $A^N v$ becomes as large as you please.

We make some definitions and put some of these properties together, for a linear dynamical system which evolves by the rule $v \mapsto Av$. First observe that 0 is always a fixed point of the system: $A(0) = 0$ no matter what A is. We say 0 is an **unstable** fixed point if we can find very small vectors w such that $|A^N(w)| \to \infty$ as $N \to \infty$. This means that, even though we start very close to our fixed point 0, we eventually end up very far away from it by applying A over and over and over. We say 0 is a **stable** fixed point of A if there is a positive number M such that

$|A^N(w)| \leq M|w|$ for all w near 0 and for all $N = 1, 2, 3, 4, \ldots$. This means that if we start near 0, then we stay near 0 forever. Finally, we say 0 is and **asymptotically stable** fixed point of A if for w close enough to 0 we have $A^N(w) \to 0$ as $N \to \infty$. This means that if we start close enough to our fixed point 0 then we actually collapse in to it.

Theorem 5 *Let A be an $n \times n$ matrix and consider the dynamical system given by $v \mapsto A(v)$. If A has an eigenvalue λ with $|\lambda| > 1$ then 0 is unstable. On the other hand, 0 is asymptotically stable if and only if $|\lambda| < 1$ for all eigenvalues λ of A.*