15. Eigenvalues, Eigenvectors

Matrix of a Linear Transformation Consider a linear transformation $L : \mathbb{R}^2 \to \mathbb{R}^2$. Suppose we know that $L\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}a\\c\end{pmatrix}$ and $L\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}b\\d\end{pmatrix}$. Then because of linearity, we can determine what L does to any vector $\begin{pmatrix}x\\y\end{pmatrix}$:

$$L\begin{pmatrix}x\\y\end{pmatrix} = L(x\begin{pmatrix}1\\0\end{pmatrix}+y\begin{pmatrix}0\\1\end{pmatrix}) = xL\begin{pmatrix}1\\0\end{pmatrix}+yL\begin{pmatrix}0\\1\end{pmatrix} = x\begin{pmatrix}a\\c\end{pmatrix}+y\begin{pmatrix}b\\d\end{pmatrix} = \begin{pmatrix}ax+by\\cx+dy\end{pmatrix}.$$

Now notice that for any vector $\begin{pmatrix}x\\y\end{pmatrix}$, we have

$$\begin{pmatrix}a & b\\c & d\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}ax+by\\cx+dy\end{pmatrix} = L\begin{pmatrix}x\\y\end{pmatrix}.$$

Then the matrix $\begin{pmatrix}a & b\\c & d\end{pmatrix}$ acts by matrix multiplication in the same way that L does. Call this matrix the *matrix of* L in the *"basis"* $\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}$.

Since every linear function from $\mathbb{R}^2 \to \mathbb{R}^2$ can be given a matrix in this way, we see that every such linear function has a matrix in the basis $\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}$. We will revisit this idea, and develop the notion of a basis further, and learn about how to make a matrix for an arbitrary linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ in an arbitrary basis.

Invariant Directions

Consider the linear transformation L such that $L\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}-4\\-10\end{pmatrix}$ and $L\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}3\\7\end{pmatrix}$, so that the matrix of L is $\begin{pmatrix}-4 & 3\\-10 & 7\end{pmatrix}$. Recall that a vector is a direction and a magnitude; L applied to $\begin{pmatrix}1\\0\end{pmatrix}$ or $\begin{pmatrix}0\\1\end{pmatrix}$ changes both the direction and the magnitude of the vectors given to it.

Notice that $L\begin{pmatrix}3\\5\end{pmatrix} = \begin{pmatrix}-4\cdot 3+3\cdot 5\\-10\cdot 3+7\cdot 5\end{pmatrix} = \begin{pmatrix}3\\5\end{pmatrix}$. Then L fixes both the magnitude and direction of the vector $v_1 = \begin{pmatrix}3\\5\end{pmatrix}$. Now, notice that any vector with the same direction as v_1 can be written as cv_1 for some constant c. Then $L(cv_1) = cL(v_1) = cv_1$, so L fixes every vector pointing in the same direction as v_1.

Also notice that $L\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}-4\cdot 1+3\cdot 2\\-10\cdot 1+7\cdot 2\end{pmatrix} = \begin{pmatrix}2\\4\end{pmatrix} = 2\begin{pmatrix}1\\2\end{pmatrix}$. Then L fixes the direction of the vector $v_2 = \begin{pmatrix}1\\2\end{pmatrix}$ but stretches v_2 by a factor of 2. Now notice that for any constant c, $L(cv_2) = cL(v_2) = 2cv_2$. Then L stretches every vector pointing in the same direction as v_2 by a factor of 2.

In short, given a linear transformation L it is sometimes possible to find a vector $v \neq 0$ and constant $\lambda \neq 0$ such that

$$L(v) = \lambda v$$

We call the direction of the vector v an *invariant direction*. In fact, any vector pointing in the same direction also satisfies the equation: $L(cv) = cL(v) = \lambda cv$. The vector v is called an *eigenvector* of L, and λ is an *eigenvalue*. Since the direction is all we really care about here, then any other vector cv (so long as $c \neq 0$) is an equally good choice of eigenvector.

Returning to our example of the linear transformation L with matrix $\begin{pmatrix}-4 & 3\\-10 & 7\end{pmatrix}$, we have seen that L enjoys the property of having two invariant directions, represented by eigenvectors v_1 and v_2 with eigenvalues 1 and 2, respectively.

It would be very convenient if I could write any vector w as a linear combination of v_1 and v_2. Suppose $w = rv_1 + sv_2$ for some constants r and s. Then:

$$L(w) = L(rv_1 + sv_2) = rL(v_1) + sL(v_2) = rv_1 + 2sv_2.$$

Now L just multiplies the number r by 1 and the number s by 2. If we could write this as a matrix, it would look like:

$$\begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}\begin{pmatrix}r\\s\end{pmatrix}$$

This is much slicker than the usual scenario, in which $L\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \begin{pmatrix}a\end{pmatrix} x+by$ $cx + dy$.

Here, r and s give the coordinates of w in terms of the vectors v_1 and v_2. In the previous example, we multiplied the vector by the matrix L and came up with a complicated expression. In these coordinates, we can see that L is a very simple *diagonal matrix*, whose diagonal entries are exactly the *eigenvalues* of L.

This process is called *diagonalization*, and it can make complicated linear systems much easier to analyze.

Now that we've seen what eigenvalues and eigenvectors are, there are a number of questions that need to be answered.

- How do we find eigenvectors and their eigenvalues?
- How many eigenvalues and (independent) eigenvectors does a given linear transformation have?
- When can a linear transformation be diagonalized?

We'll start by trying to find the eigenvectors for a linear transformation.

Example Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ such that $L(x, y) = (2x + 2y, 16x + 6y)$. First, we can find the matrix of L:

$$\begin{pmatrix}x\\ y\end{pmatrix} \stackrel{L}{\mapsto} \begin{pmatrix}2 & 2\\ 16 & 6\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}.$$

We want to find an invariant direction $v = \begin{pmatrix}x\\ y\end{pmatrix}$ such that

$$L(v) = \lambda v$$

or, in matrix notation,

$$\begin{aligned}
\begin{pmatrix}2 & 2\\ 16 & 6\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} &= \lambda\begin{pmatrix}x\\ y\end{pmatrix}\\
\Leftrightarrow \begin{pmatrix}2 & 2\\ 16 & 6\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} &= \begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}\\
\Leftrightarrow \begin{pmatrix}2-\lambda & 2\\ 16 & 6-\lambda\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} &= \begin{pmatrix}0\\ 0\end{pmatrix}.
\end{aligned}$$

This is a homogeneous system, so it only has solutions when the matrix $\begin{pmatrix} 2-\lambda & 2 \\ 16 & 6-\lambda \end{pmatrix}$ is singular. In other words,

$$\begin{aligned} \det \begin{pmatrix} 2-\lambda & 2 \\ 16 & 6-\lambda \end{pmatrix} &= 0 \\ \Leftrightarrow (2-\lambda)(6-\lambda) - 32 &= 0 \\ \Leftrightarrow \lambda^2 - 8\lambda - 20 &= 0 \\ \Leftrightarrow (\lambda - 10)(\lambda + 2) &= 0 \end{aligned}$$

For any square $n \times n$ matrix M, the polynomial in λ given by $\det(\lambda I - M) = (-1)^n \det(M - \lambda I)$ is called the *characteristic polynomial* of M, and its roots are the eigenvalues of M.

In this case, we see that L has two eigenvalues, $\lambda_1 = 10$ and $\lambda_2 = -2$. To find the eigenvectors, we need to deal with these two cases separately.

To do so, we solve the linear system $\begin{pmatrix} 2-\lambda & 2 \\ 16 & 6-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ with the particular eigenvalue λ plugged in to the matrix.

$\underline{\lambda = 10}$: We solve the linear system

$$\begin{pmatrix} -8 & 2 \\ 16 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Both equations say that $y = 4x$, so any vector $\begin{pmatrix} x \\ 4x \end{pmatrix}$ will do. Since we only need the direction of the eigenvector, we can pick a value for x. Setting $x = 1$ is convenient, and gives the eigenvector $v_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$.

$\underline{\lambda = -2}$: We solve the linear system

$$\begin{pmatrix} 4 & 2 \\ 16 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Here again both equations agree, because we chose λ to make the system singular. We see that $y = -2x$ works, so we can choose $v_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

In short, our process was the following:

- Find the characteristic polynomial of the matrix M for L, given by $\det(\lambda I - M)$.
- Find the roots of the characteristic polynomial; these are the eigenvalues of L.
- For each eigenvalue λ_i, solve the linear system $(\lambda_i I - M)x = 0$ to obtain an eigenvector x associated to λ_i.

References

- Hefferon, Chapter Three, Section III.1: Representing Linear Maps with Matrices
- Hefferon, Chapter Five, Section II.3: Eigenvalues and Eigenvectors

Wikipedia:

- Eigen*
- Characteristic Polynomial
- Linear Transformations (and matrices thereof)

Review Questions

1. Consider $L : \mathbb{R}^2 \to \mathbb{R}^2$ with $L(x, y) = (x \cos\theta + y \sin\theta, -x \sin\theta + y \cos\theta)$.

   i. Write the matrix of L on the basis $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$.

   ii. When $\theta \neq 0$, explain how L acts on the plane. Draw a picture.

   iii. Do you expect L to have invariant directions?

   iv. Try to find eigenvalues for L by solving the equation

   $$L(v) = \lambda v.$$

   v. Does L have real eigenvalues? If not, are there complex eigenvalues for L, assuming that $i = \sqrt{-1}$ exists?

2. Let $M = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. Find all eigenvalues of M. Does M have two independent eigenvectors? Can M be diagonalized?

3. Let L be the linear transformation $L : \mathbb{R}^3 \to \mathbb{R}^3$ given by $L(x, y, z) = (x+y, y+z, x+z)$. Let e_i be the vector with a one in the ith position and zeros in all other positions.

   i. Find Le_i for each i.

   ii. Given a matrix $M = \begin{pmatrix} m_1^1 & m_2^1 & m_3^1 \\ m_1^2 & m_2^2 & m_3^2 \\ m_1^3 & m_2^3 & m_3^3 \end{pmatrix}$, what can you say about Me_i for each i?

   iii. Find a 3×3 matrix M representing L. Choose three non-trivial vectors pointing in different directions and show that $Mv = Lv$ for each of your choices.