CHAPTER 10 EIGENVALUES AND EIGENVECTORS

Definition 10.0.3: Eigenvalues and Eigenvectors

For a square matrix $\mathbf{A}_{n\times n}$, a scalar λ is called an **eigenvalue** of **A** if there is a nonzero vector **x** such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

Such a vector, **x** is called an **eigenvector** of **A** corresponding to the eigenvalue λ. We sometimes refer to the pair $(\lambda, \mathbf{x})$ as an **eigenpair**.

Eigenvalues and eigenvectors have numerous applications throughout mathematics, statistics and other fields. First, we must get a handle on the definition which we will do through some examples.

Example 10.0.3: Eigenvalues and Eigenvectors

Determine whether $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector of $\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ and if so, find the corresponding eigenvalue.
To determine whether **x** is an eigenvector, we want to compute **Ax** and observe whether the result is a multiple of **x**. If this is the case, then the multiplication factor is the corresponding eigenvalue:

$$\mathbf{A}\mathbf{x} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

From this it follows that **x** *is* an eigenvector of **A** and the corresponding eigenvalue is $\lambda = 4$.

Is the vector $\mathbf{y} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ an eigenvector?

$$\mathbf{Ay} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ 8 \end{pmatrix} = 4 \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 4\mathbf{y}$$

Yes, it is and it corresponds to the *same* eigenvalue, $\lambda = 4$

Example 10.0.3 shows a very important property of eigenvalue-eigenvector pairs. If $(\lambda, \mathbf{x})$ is an eigenpair then any scalar multiple of $\mathbf{x}$ is also an eigenvector corresponding to λ. To see this, let $(\lambda, \mathbf{x})$ be an eigenpair for a matrix $\mathbf{A}$ (which means that $\mathbf{Ax} = \lambda\mathbf{x}$) and let $\mathbf{y} = \alpha\mathbf{x}$ be any scalar multiple of $\mathbf{x}$. Then we have,

$$\mathbf{Ay} = \mathbf{A}(\alpha\mathbf{x}) = \alpha(\mathbf{Ax}) = \alpha(\lambda\mathbf{x}) = \lambda(\alpha\mathbf{x}) = \lambda\mathbf{y}$$

which shows that $\mathbf{y}$ (or any scalar multiple of $\mathbf{x}$) is also an eigenvector associated with the eigenvalue λ.

Thus, for each eigenvalue we have infinitely many eigenvectors. In the preceding example, the eigenvectors associated with $\lambda = 4$ will be scalar multiples of $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. You may recall from Chapter 6 that the set of all scalar multiples of $\mathbf{x}$ is denoted $span(\mathbf{x})$. The $span(\mathbf{x})$ in this example represents the **eigenspace** of λ. *Note: when using software to compute eigenvectors, it is standard practice for the software to provide the normalized/unit eigenvector.*

In some situations, an eigenvalue can have multiple eigenvectors which are linearly independent. The number of linearly independent eigenvectors associated with an eigenvalue is called the **geometric multiplicity** of the eigenvalue. Example 10.0.4 clarifies this concept.

Example 10.0.4: Geometric Multiplicity

Consider the matrix $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$. It should be straightforward to see that $\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are both eigenvectors corresponding to the eigenvalue $\lambda = 3$. $\mathbf{x}_1$ and $\mathbf{x}_2$ are linearly independent, therefore the geometric multiplicity of $\lambda = 3$ is 2.

What happens if we take a linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$? Is that also an eigenvector? Consider $\mathbf{y} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2\mathbf{x}_1 + 3\mathbf{x}_2$. Then

$$\mathbf{Ay} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \end{pmatrix} = 3 \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 3\mathbf{y}$$

shows that $\mathbf{y}$ is also an eigenvector associated with $\lambda = 3$.
The **eigenspace** corresponding to $\lambda = 3$ is the set of all linear combinations of $\mathbf{x}_1$ and $\mathbf{x}_2$, i.e. the $span(\mathbf{x}_1, \mathbf{x}_2)$.

We can generalize the result that we saw in Example 10.0.4 for any square matrix and any geometric multiplicity. Let $\mathbf{A}_{n\times n}$ have an eigenvalue λ with geometric multiplicity k. This means there are k linearly independent eigenvectors, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ such that $\mathbf{Ax}_i = \lambda\mathbf{x}_i$ for each eigenvector $\mathbf{x}_i$. Now if we let $\mathbf{y}$ be a vector in the $span(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ then $\mathbf{y}$ is some linear combination of the $\mathbf{x}_i$'s:

$$\mathbf{y} = \alpha_1\mathbf{x}_2 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k$$

Observe what happens when we multiply $\mathbf{y}$ by $\mathbf{A}$:

$$\begin{aligned}
\mathbf{Ay} &= \mathbf{A}(\alpha_1\mathbf{x}_2 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k) \\
&= \alpha_1(\mathbf{Ax}_1) + \alpha_2(\mathbf{Ax}_2) + \cdots + \alpha_k(\mathbf{Ax}_k) \\
&= \alpha_1(\lambda\mathbf{x}_1) + \alpha_2(\lambda\mathbf{x}_2) + \cdots + \alpha_k(\lambda\mathbf{x}_k) \\
&= \lambda(\alpha_1\mathbf{x}_2 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k) \\
&= \lambda\mathbf{y}
\end{aligned}$$

which shows that $\mathbf{y}$ (or any vector in the $span(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$) is an eigenvector of $\mathbf{A}$ corresponding to λ.

This proof allows us to formally define the concept of an eigenspace.

Definition 10.0.4: Eigenspace

Let $\mathbf{A}$ be a square matrix and let λ be an eigenvalue of $\mathbf{A}$. The set of all eigenvectors corresponding to λ, together with the zero vector, is called the **eigenspace** of λ. The number of basis vectors required to form the eigenspace is called the **geometric multiplicity** of λ.

Now, let's attempt the eigenvalue problem from the other side. Given an eigenvalue, we will find the corresponding eigenspace in Example 10.0.5.

Example 10.0.5: Eigenvalues and Eigenvectors

Show that $\lambda = 5$ is an eigenvalue of $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ and determine the eigenspace of $\lambda = 5$.

Attempting the problem from this angle requires slightly more work. We want to find a vector $\mathbf{x}$ such that $\mathbf{Ax} = 5\mathbf{x}$. Setting this up, we have:

$$\mathbf{Ax} = 5\mathbf{x}.$$

What we want to do is move both terms to one side and factor out the vector x. In order to do this, we must use an identity matrix, otherwise the equation wouldn't make sense (we'd be subtracting a constant from a matrix).

$$\begin{aligned} \mathbf{Ax} - 5\mathbf{x} &= \mathbf{0} \\ (\mathbf{A} - 5\mathbf{I})\mathbf{x} &= \mathbf{0} \\ \left(\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} - \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} -4 & 2 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

Clearly, the matrix $\mathbf{A} - \lambda\mathbf{I}$ is singular (i.e. does not have linearly independent rows/columns). This will always be the case by the definition $\mathbf{Ax} = \lambda\mathbf{x}$, and is often used as an alternative definition.
In order to solve this homogeneous system of equations, we use Gaussian elimination:

$$\left(\begin{array}{cc|c} -4 & 2 & 0 \\ 4 & -2 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right)$$

This implies that any vector $\mathbf{x}$ for which $x_1 - \frac{1}{2}\mathbf{x}_2 = 0$ satisfies the eigenvector equation. We can pick any such vector, for example $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, and say that the eigenspace of $\lambda = 5$ is

$$span \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$$

If we didn't know either an eigenvalue or eigenvector of $\mathbf{A}$ and instead wanted to find both, we would first find eigenvalues by determining all possible λ such that $\mathbf{A} - \lambda\mathbf{I}$ is singular and then find the associated eigenvectors. There are some tricks which allow us to do this by hand for 2×2 and 3×3 matrices, but after that the computation time is unworthy of the effort. Now that we have a good understanding of how to interpret eigenvalues and eigenvectors algebraically, let's take a look at some of the things that they can do, starting with one important fact.

Theorem 10.0.2: Eigenvalues and the Trace of a Matrix

Let $\mathbf{A}$ be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then the sum of the eigenvalues is equal to the trace of the matrix (recall that the trace of a matrix is the sum of its diagonal elements).

$$Trace(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i.$$

Example 10.0.6: Trace of Covariance Matrix

Suppose that we had a collection of n observations on p variables, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p$. After centering the data to have zero mean, we can compute the sample variances as:

$$var(\mathbf{x}_i) = \frac{1}{n-1}\mathbf{x}_i^T\mathbf{x}_i = \|\mathbf{x}_i\|^2$$

These variances form the diagonal elements of the sample covariance matrix,

$$\mathbf{\Sigma} = \frac{1}{n-1}\mathbf{X}^T\mathbf{X}$$

Thus, the total variance of this data is

$$\frac{1}{n-1}\sum_{i=1}^{n}\|\mathbf{x}_i\|^2 = Trace(\mathbf{\Sigma}) = \sum_{i=1}^{n}\lambda_i.$$

In other words, the sum of the eigenvalues of a covariance matrix provides the total variance in the variables $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

10.1 Diagonalization

Let's take another look at Example 10.0.5. We already showed that $\lambda_1 = 5$ and $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is an eigenpair for the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$. You may verify that $\lambda_2 = -1$ and $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is another eigenpair. Suppose we create a matrix of eigenvectors:

$$\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2) = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$$

and a diagonal matrix containing the corresponding eigenvalues:

$$\mathbf{D} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$$

Then it is easy to verify that $\mathbf{AV} = \mathbf{VD}$:

$$\begin{aligned} \mathbf{AV} &= \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 5 & -1 \\ 10 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \mathbf{VD} \end{aligned}$$

If the columns of $\mathbf{V}$ are linearly independent, which they are in this case, we can write:

$$\mathbf{V}^{-1}\mathbf{AV} = \mathbf{D}$$

What we have just done is develop a way to transform a matrix $\mathbf{A}$ into a diagonal matrix $\mathbf{D}$. This is known as **diagonalization**.

Definition 10.1.1: Diagonalizable

An $n \times n$ matrix $\mathbf{A}$ is said to be **diagonalizable** if there exists an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$$

This is possible if and only if the matrix $\mathbf{A}$ has n linearly independent eigenvectors (known as a complete set of eigenvectors). The matrix $\mathbf{P}$ is then the matrix of eigenvectors and the matrix $\mathbf{D}$ contains the corresponding eigenvalues on the diagonal.

Determining whether or not a matrix $\mathbf{A}_{n \times n}$ is diagonalizable is a little tricky. Having $rank(\mathbf{A}) = n$ is *not* a sufficient condition for having n linearly independent eigenvectors. The following matrix stands as a counter example:

$$\mathbf{A} = \begin{pmatrix} -3 & 1 & -3 \\ 20 & 3 & 10 \\ 2 & -2 & 4 \end{pmatrix}$$

This matrix has full rank but only two linearly independent eigenvectors. Fortunately, for our primary application of diagonalization, we will be dealing with a symmetric matrix, which can always be diagonalized. In fact, symmetric matrices have an additional property which makes this diagonalization particularly nice, as we will see in Chapter 11.

10.2 Geometric Interpretation of Eigenvalues and Eigenvectors

Since any scalar multiple of an eigenvector is still an eigenvector, let's consider for the present discussion unit eigenvectors $\mathbf{x}$ of a square matrix $\mathbf{A}$ - those with length $\|\mathbf{x}\| = 1$. By the definition, we know that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

We know that geometrically, if we multiply $\mathbf{x}$ by $\mathbf{A}$, the resulting vector points in the same direction as $\mathbf{x}$. Geometrically, it turns out that multiplying the unit circle or unit sphere by a matrix $\mathbf{A}$ carves out an ellipse, or an ellipsoid. We can see eigenvectors visually by watching how multiplication by a matrix $\mathbf{A}$ changes the unit vectors. Figure 10.1 illustrates this. The blue arrows represent (a sampling of) the unit circle, all vectors $\mathbf{x}$ for which $\|\mathbf{x}\| = 1$. The red arrows represent the image of the blue arrows after multiplication by $\mathbf{A}$, or $\mathbf{A}\mathbf{x}$ for each vector $\mathbf{x}$. We can see how almost every vector changes direction when multiplied by $\mathbf{A}$, except the eigenvector directions which are marked in black. Such a picture provides a nice geometrical interpretation of eigenvectors for a general matrix, but we will see in Chapter 11 just how powerful these eigenvector directions are when we look at symmetric matrix.

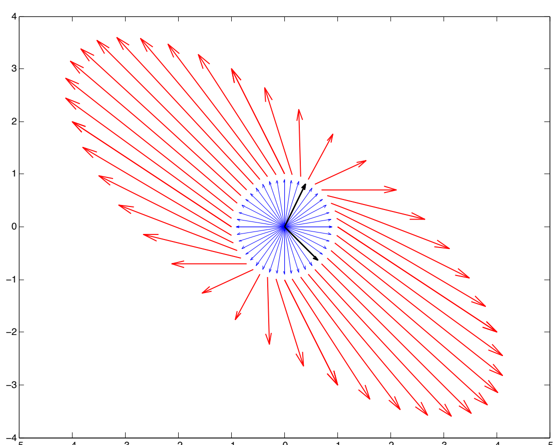

Figure 10.1: Visualizing eigenvectors (in black) using the image (in red) of the unit sphere (in blue) after multiplication by $\mathbf{A}$.

Exercises

1. Show that **v** is an eigenvector of **A** and find the corresponding eigenvalue:

   a. $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \quad \mathbf{v} = \begin{pmatrix} 3 \\ -3 \end{pmatrix}$

   b. $\mathbf{A} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix} \quad \mathbf{v} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

   c. $\mathbf{A} = \begin{pmatrix} 4 & -2 \\ 5 & -7 \end{pmatrix} \quad \mathbf{v} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

2. Show that λ is an eigenvalue of **A** and list two eigenvectors corresponding to this eigenvalue:

   a. $\mathbf{A} = \begin{pmatrix} 0 & 4 \\ -1 & 5 \end{pmatrix} \quad \lambda = 4$

   b. $\mathbf{A} = \begin{pmatrix} 0 & 4 \\ -1 & 5 \end{pmatrix} \quad \lambda = 1$

3. Based on the eigenvectors you found in exercises 2, can the matrix **A** be diagonalized? Why or why not? If diagonalization is possible, explain how it would be done.

4. Can a rectangular matrix have eigenvalues/eigenvectors?