1 Eigenvector and Eigenvalues (5.1, 5.2)

Fundamentally, we want to solve

$$Av = \lambda v, \tag{1}$$

where

- A is a matrix either in $\mathbb{R}^{n\times n}$ or $\mathbb{C}^{n\times n}$, and
- $\lambda \in \mathbb{C}$ (the *eigenvalue*), and
- $v \in \mathbb{C}^n$ (the *eigenvector*).

The pair (λ, v) is called an **eigenpair** of A. Whereas each eigenvector has a unique eigenvalue associated with it, each eigenvalue is associated with many eigenvectors. For example, if v is an eigenvector of A associated with the eigenvalue λ, then every nonzero multiple of v is also an eigenvector of A associated with λ. Now, note that

$$Av = \lambda v \implies Av = \lambda I v \implies \mathbf{0} = Av - \lambda I v \implies \mathbf{0} = (A - \lambda I)v,$$

where I is the $n \times n$ identity matrix. In any case, if v is an eigenvector with eigenvalue λ, then v is a nonzero solution of the homogeneous matrix equation

$$(A - \lambda I)v = \mathbf{0}.$$

Therefore, the matrix is singular[1] and $\det(A - \lambda I) = 0$. This establishes the following theorem.

Theorem 1.1

λ is an eigenvalue of A if and only if

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

This equation is known as the **characteristic equation** of A.

Remark: Although this is useful theoretically, this equation has little value for actual computations, as we will see later.

We note that $\det(A - \lambda I)$ is a polynomial in λ of degree n; it is called the **characteristic polynomial** of A. Therefore, the eigenvalues, λ, are the roots of this polynomial. There are exactly n roots, not all real (i.e., some complex), and thus n eigenvalues.

(Exercise.) Find the characteristic polynomial of

$$B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

We have

$$\begin{aligned} \det(B - \lambda I) &= \det\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} 1-\lambda & 2 \\ 3 & 4-\lambda \end{bmatrix}\right) \\ &= (1-\lambda)(4-\lambda) - 3(2) \\ &= 4 - \lambda - 4\lambda + \lambda^2 - 6 \\ &= \lambda^2 - 5\lambda - 2. \end{aligned}$$

[1] Not invertible.

1.1 Getting Eigenvectors

We can get the eigenvectors by solving

$$(A - \lambda I)v = \mathbf{0}, \tag{2}$$

where λ are the eigenvalues that we found by solving (1.1). Note that

- Eigenvalues are the roots of the polynomials of (1.1).
- Roots can be complex, thus eigenvalues can also be complex *even* when A is real.

(Example.) Consider

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

- To find the eigenvalues, we have

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

  Then, we can solve

$$\lambda^2 + 1 = 0,$$

  giving us $\lambda = \pm i$, which is *complex*.

- To find the eigenvectors, we use the formula (2) for each eigenvalue.
  - For $\lambda = i$, we have

$$\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix}\right) v = \mathbf{0} \implies \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} v = \mathbf{0} \implies v = \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

  - For $\lambda = -i$, we have

$$\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} -i & 0 \\ 0 & -i \end{bmatrix}\right) v = \mathbf{0} \implies \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} v = \mathbf{0} \implies v = \begin{bmatrix} -i \\ 1 \end{bmatrix}.$$

1.2 Eigenvectors & Eigenvalues for Large Matrices

Suppose we have a matrix of size 200×200. The characteristic equation will be a polynomial of degree 200. This means that there are 200 roots and thus 200 eigenvalues. Now, we should note that there is no formula to find the roots when $r > 4$ (we have the quadratic equation when $r = 2$, for example), so this process is tedious. Additionally, numerically finding the roots is ill-conditioned[2].

So, how do we find the eigenvalues numerically? As implied earlier, the characteristic equation is not useful for finding eigenvalues numerically, so we will perform iterative methods to find eigenvalues.

1.2.1 QR Iteration

Before we begin this section, let's note a theorem.

Theorem 1.2

Let $T \in \mathbb{C}^{n\times n}$ be a (lower- or upper-) triangular matrix. Then, the eigenvalues of T are the main-diagonal entries $t_{11}, \ldots, t_{nn}$.

We will construct a sequence of matrices with the same eigenvalues that converge to a triangular matrix. The reason why we choose a triangular matrix is because its eigenvalues are on the diagonal.

[2]Small perturbations to the coefficients lead to large changes in the roots.

1.3 Semisimple Matrices

Definition 1.1: Semisimple Matrix

A matrix $A \in \mathbb{R}^{n\times n}$ is called **semisimple** if it has n linearly independent eigenvectors. Otherwise, it is called **defective.**

(Example.) Consider

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Its characteristic polynomial is given by

$$\det(I - \lambda I) = (1-\lambda)^2.$$

Solving $(1-\lambda)^2 = 0$ yields $\lambda_1 = \lambda_2 = 1$. So, we find $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, which are are linearly independent, which means that I is semisimple.

Remark: Any vector, except $\mathbf{0}$, can be an eigenvector.

(Example.) Consider

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}.$$

Its characteristic polynomial is given by

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} 1-\lambda & 3 \\ 0 & 1-\lambda \end{bmatrix}\right) = (1-\lambda)^2.$$

Solving $(1-\lambda)^2 = 0$ yields $\lambda_1 = \lambda_2 = 1$. To get the eigenvectors, we solve

$$(A - \lambda I)v = \mathbf{0},$$

so in particular

$$\left(\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} - 1 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) v = \mathbf{0} \implies \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \mathbf{0} \implies \begin{cases} 0v_1 + 3v_2 = 0 \\ 0v_1 + 0v_2 = 0 \end{cases}.$$

We find that $3v_2 = 0 \implies v_2 = 0$ and v_1 can be any arbitrary value *except* 0 (because eigenvectors cannot be $\mathbf{0}$). In particular, every vector of the form

$$v = \begin{bmatrix} v_1 \\ 0 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

is an eigenvector. However, this implies that there is only one linearly independent eigenvector hence A is defective.

1.3.1 Reminders

Recall that

- if v is an eigenvector of A with respect to λ, then cv for some $c \in \mathbb{R} \setminus \{0\}$ is also an eigenvector of A with respect to λ. In particularm

$$A(cv) = c(Av) = c(\lambda v) = \lambda(cv) \implies A(cv) = \lambda(cv).$$

- if v is an eigenvector of A with respect to λ, then v is also an eigenvector of A^{-1} with respect to λ^{-1}.

$$Av = \lambda v \implies A^{-1}(Av) = A^{-1}(\lambda v) \implies v = \lambda A^{-1} v \implies \lambda^{-1} v = A^{-1} v.$$

This brings us to another theorem.

Theorem 1.3

If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of A, then the associated eigenvectors $v_1, v_2, \ldots, v_k$ are linearly independent.