Chapter 8 Eigenvalues, Eigenvectors, and Diagonalization

- Recall that a matrix transformation is, a function f from R^m to R^n determined by an $m \times n$ matrix **A** so that f(**p**) = **Ap**, for **p** in R^m.

The length of input vector **p** need not be the same as the length of the image vector **Ap**; that is ||**p**|| need not equal ||**Ap**||.

Since a vector **p** in R^n has two basic properties, length and direction, it is reasonable to investigate whether **p** and **Ap** are ever parallel for a given matrix **A**. (Recall that parallel means that one vector is a scalar multiple of the other; **Ap** = *(scalar)* ∗**p .**)

We note that A0 = 0, hence we need consider only nonzero vectors p.

The vectors p such that **Ap** and **p** are parallel arise in a surprising number of places, including **vibrations, aerodynamics, elasticity, nuclear physics, mechanics, chemical engineering, biology, digital transmission of images**, etc.

Here we consider only square matrices. We begin with 2 × 2 and 3 × 3 matrices, but our techniques also apply to larger square matrices.

++

Example Let $\mathbf{A} = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\mathbf{q} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Then $\mathbf{Ap} = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3\mathbf{p}$ and

$$\mathbf{Aq} = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix} = -3\begin{bmatrix} -1 \\ 1 \end{bmatrix} = -3\mathbf{q}$$

+++

Example Let $\mathbf{A} = \mathbf{I_n}$ (the n × n identity), then for every n-vector **x**, **Ax** = 1**x**.

+++

Example Let $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Let $\mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix}$. Determine all vectors **p** so that **Ap** is parallel to **p**.

Strategy: Compute **Ap** and set it equal to scalar λ times **p**. Use equality of vectors to determine λ and the components of vector **p**.

$\mathbf{Ap} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ -\mathbf{y} \end{bmatrix} = \lambda\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \lambda\mathbf{x} \\ \lambda\mathbf{y} \end{bmatrix}$ then setting corresponding entries equal to zero we have two cases to investigate:

Case1: **x = λx**, so for **x** not 0 we have **λ =1**; then for **–y = λy** we have **–y = y** so that **y = 0**.

So **λ = 1** and $\mathbf{p} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}$ where **x** ≠ 0, since we do not consider the zero vector a solution.

Thus $\mathbf{Ap} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} = 1\begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}$

Case 2: **-y = λy**, so for y not 0 we have **λ = -1;** then for **x = λx** we have **x = -x**, so that **x = 0**.

So **λ = -1** and $\mathbf{p} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}$ where **y** ≠ 0, since we do not consider the zero vector a solution.

Thus $\mathbf{Ap} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} = -1\begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}$

> Summary: the set of vectors p so that **Ap** =λ**p**, are all vectors in R^2 of the form $\left\{\begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}, \mathbf{x} \neq 0 \text{ and } \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}, \mathbf{y} \neq 0\right\}$. Note that only particular values of scalar λ can be chosen.

++

<u>Example</u> Let $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}$. Let $\mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix}$. Determine all vectors **p** so that **Ap** is parallel to **p**.

$\mathbf{Ap} = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} 2\mathbf{x} - \mathbf{y} \\ 3\mathbf{y} \end{bmatrix} = \lambda\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \lambda\mathbf{x} \\ \lambda\mathbf{y} \end{bmatrix}$ then setting corresponding entries equal to zero we have two cases to investigate:

<u>Case 1:</u> 2**x - y = λx**, so for **x** not 0 we have **y = 0** and **λ = 2;** then for **3y = 2y** we also have **y = 0.**

So **λ = 2** and $\mathbf{p} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}$ where **x** ≠ 0, since we do not consider the zero vector a solution.

Thus $\mathbf{Ap} = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} = 2\begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}$

<u>Case 2:</u> **3y = λy**, so for **y** not 0 we have **λ = 3;** then for 2**x - y = 3x** we have **x = -y**.

So **λ = 3** and $\mathbf{p} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} -\mathbf{y} \\ \mathbf{y} \end{bmatrix}$ where **y** ≠ 0, since we do not consider the zero vector a solution.

Thus $\mathbf{Ap} = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} -\mathbf{y} \\ \mathbf{y} \end{bmatrix} = 3\begin{bmatrix} -\mathbf{y} \\ \mathbf{y} \end{bmatrix}$

> Summary: the set of vectors p so that **Ap** =λ**p**, are all vectors in R^2 of the form $\left\{\begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}, \mathbf{x} \neq 0 \text{ and } \begin{bmatrix} -\mathbf{y} \\ \mathbf{y} \end{bmatrix}, \mathbf{y} \neq 0\right\}$. Note that only particular values of scalar λ can be chosen.

Definition Let **A** be an $n \times n$ matrix. The scalar λ is called an **eigenvalue** of matrix **A** if there exists an $n \times 1$ vector **x**, $\mathbf{x} \neq \mathbf{0}$, such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}. \qquad (1)$$

Every nonzero vector **x** satisfying (1) is called an **eigenvector of A associated with eigenvalue** λ.

Equation (1) is commonly called the **eigen equation** and its solution is a goal of this chapter.

Sometimes we will use the term **eigenpair** to mean **an eigenvalue and an associated eigenvector**. This emphasizes that for a given matrix **A** a solution of (1) requires both a scalar, an eigenvalue λ, and a nonzero vector, an eigenvector **x**.

In the eigen equation in (1) there is no restriction on the entries of matrix **A**, the type of scalar λ, or the entries of vector **x**, other than they cannot all be zero. It is possible that an eigen values and eigenvectors of a matrix with real entries can involve complex values.

Computing Eigen Information for Small Matrices

The eigen equation can be rearranged as follows:

$$\mathbf{Ax} = \lambda\mathbf{x} \Leftrightarrow \mathbf{Ax} = \lambda\mathbf{I_n x} \Leftrightarrow \mathbf{Ax} - \lambda\mathbf{I_n x} = \mathbf{0} \Leftrightarrow (\mathbf{A} - \lambda\mathbf{I_n})\mathbf{x} = \mathbf{0} \qquad (1)$$

The matrix equation $(\mathbf{A} - \lambda\mathbf{I_n})\mathbf{x} = \mathbf{0}$ is a homogeneous linear system with coefficient matrix $\mathbf{A} - \lambda\mathbf{I_n}$. Since an eigenvector **x** cannot be the zero vector, this means we seek a nontrivial solution to the linear system $(\mathbf{A} - \lambda\mathbf{I_n})\mathbf{x} = \mathbf{0}$. Thus $\mathbf{ns}(\mathbf{A} - \lambda\mathbf{I_n}) \neq \mathbf{0}$ or equivalently $\mathbf{rref}(\mathbf{A} - \lambda\mathbf{I_n})$ must contain a zero row. It follows that matrix $\mathbf{A} - \lambda\mathbf{I_n}$ must be singular, so from Chapter 2,

$$\mathbf{det}(\mathbf{A} - \lambda\mathbf{I_n}) = 0. \qquad (2)$$

Equation (2) is called the **characteristic equation** of matrix **A** and solving it for λ gives us the eigenvalues of **A**. Because the determinant is a linear combination of particular products of entries of the matrix, the characteristic equation is really a polynomial equation of degree n. We call

$$\mathbf{c}(\lambda) = \mathbf{det}(\mathbf{A} - \lambda\mathbf{I_n}) \qquad (3)$$

the **characteristic polynomial** of matrix **A**. The eigenvalues are the solutions of (2) or equivalently **the roots of the characteristic polynomial** (3). Once we have the n eigenvalues of **A**, λ_1, λ_2, ..., λ_n, the corresponding eigenvectors are **nontrivial solutions of the homogeneous linear systems**

$$(\mathbf{A} - \lambda_i\mathbf{I_n})\mathbf{x} = \mathbf{0} \quad \text{for } i = 1,2, ..., n. \qquad (4)$$

We summarize the computational approach for determining eigenpairs (λ,**x**) as a two-step procedure:

Step I. To find the eigenvalues of **A** compute the roots of the characteristic equation $\mathbf{det}(\mathbf{A} - \lambda\mathbf{I_n}) = 0$.

Step II. To find an eigenvector corresponding to an eigenvalue μ, compute a nontrivial solution to the homogeneous linear system $(\mathbf{A} - \mu\mathbf{I_n})\mathbf{x} = \mathbf{0}$.

Example 1. Let $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$. To find eigenpairs of **A** we follow the two step procedure given above.

Step I. Find the eigenvalues.

$$\mathbf{c}(\lambda) = \mathbf{det}(\mathbf{A} - \lambda\mathbf{I}_2) = \mathbf{det}\left(\begin{bmatrix} 1-\lambda & 1 \\ -2 & 4-\lambda \end{bmatrix}\right)$$

$$= (1-\lambda)(4-\lambda) + 2 = \lambda^2 - 5\lambda + 6$$

Thus the characteristic polynomial is a quadratic and the eigenvalues are the solutions of $\lambda^2 - 5\lambda + 6 = 0$. We factor the quadratic to get $(\lambda-3)(\lambda-2) = 0$ so the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 2$.

Step II. To find corresponding eigenvectors we solve Equation (4).

Case $\lambda_1 = 3$: We have that $(\mathbf{A} - 3\mathbf{I}_2)\mathbf{x} = \mathbf{0}$ has augmented matrix $\left[\begin{array}{cc|c} -2 & 1 & 0 \\ -2 & 1 & 0 \end{array}\right]$ and its rref is $\left[\begin{array}{cc|c} 1 & \frac{-1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right]$. (Verify.) Thus if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ we have $x_1 = (1/2)x_2$ so $\mathbf{x} = \begin{bmatrix} (1/2)x_2 \\ x_2 \end{bmatrix} = x_2\begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$, $x_2 \neq 0$. Choosing $x_2 = 2$, to conveniently get integer entries, gives eigenvector $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Case $\lambda_2 = 2$: We have that $(\mathbf{A} - 2\mathbf{I}_2)\mathbf{p} = \mathbf{0}$ has augmented matrix $\left[\begin{array}{cc|c} -1 & 1 & 0 \\ -2 & 2 & 0 \end{array}\right]$ and its rref is $\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$. (Verify.) Thus if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ we have $x_1 = x_2$ so $\mathbf{x} = \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} = x_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $x_2 \neq 0$. Choosing $x_2 = 1$ gives eigenvector $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

■

> Summary: For matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$ we found eigenvalue $\lambda = 3$ and an associated eigenvector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and for eigenvalue $\lambda = 2$ we found an associated eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Given that λ is an eigenvalue of **A**, then we know that matrix $\mathbf{A} - \lambda\mathbf{I}_n$ is singular and hence $\mathbf{rref}(\mathbf{A} - \lambda\mathbf{I}_n)$ will have at least one zero row. A homogeneous linear system, as in (4), whose coefficient matrix has rref with at least one zero row will have a solution set with at least one free variable. The free variables can be chosen to have any value as long as the resulting solution is not the zero vector. In Example 1 for eigenvalue $\lambda_1 = 3$, the general solution of (4) was $x_2\begin{bmatrix}1/2\\1\end{bmatrix}$. The free variable in this case could be any nonzero value. We chose $x_2 = 2$ to avoid fractions, but this is not required. If we chose $x_2 = 1/7$, then $\mathbf{x} = \begin{bmatrix}1/14\\1/7\end{bmatrix}$ is a valid eigenvector.

If $\mathbf{x}$ is an eigenvector corresponding to eigenvalue λ of $\mathbf{A}$, then so is $k\mathbf{x}$ for any scalar $k \neq 0$.
Proof: $\mathbf{Ax} = \lambda\mathbf{x}$ and $\mathbf{x} \neq \mathbf{0}$ since it is an eigenvector. Let $k \neq 0 \Rightarrow k\mathbf{x} \neq \mathbf{0}$.
Thus, $\mathbf{A}(k\mathbf{x}) = k(\mathbf{Ax}) = k(\lambda\mathbf{x}) = \lambda(k\mathbf{x})$,
which shows that $k\mathbf{x}$ is an eigenvector of $\mathbf{A}$.

Example 2. Find the eigenvalues and corresponding eigenvectors of matrix $\mathbf{A} = \frac{1}{2}\begin{bmatrix}\sqrt{2} & -\sqrt{2}\\ \sqrt{2} & \sqrt{2}\end{bmatrix}$. We follow our two step procedure.

Step I. Find the eigenvalues.

$$\mathbf{c}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}_2) = \det\left(\begin{bmatrix}\frac{\sqrt{2}}{2} - \lambda & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} - \lambda\end{bmatrix}\right)$$

$$= \left(\frac{\sqrt{2}}{2} - \lambda\right)^2 + 2 = \lambda^2 - \sqrt{2}\lambda + 1$$

Applying the quadratic formula to find the roots of $\mathbf{c}(\lambda) = 0$ we get $\lambda = \frac{\sqrt{2} \pm i\sqrt{2}}{2}$. Thus we have a pair of complex eigenvalues

$$\lambda_1 = \frac{\sqrt{2} + i\sqrt{2}}{2} \quad \text{and} \quad \lambda_2 = \frac{\sqrt{2} - i\sqrt{2}}{2}$$

even though all the entries in $\mathbf{A}$ are real.

Step II. To find corresponding eigenvectors we solve Equation (4).

Case $\lambda_1 = \frac{\sqrt{2} + i\sqrt{2}}{2}$: We have that $\left(\mathbf{A} - \frac{\sqrt{2} + i\sqrt{2}}{2}\mathbf{I}_2\right)\mathbf{p} = \mathbf{0}$ has augmented matrix $\left[\begin{array}{cc|c}-i\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0\\ \frac{\sqrt{2}}{2} & -i\frac{\sqrt{2}}{2} & 0\end{array}\right]$ and its rref is $\left[\begin{array}{cc|c}1 & -i & 0\\ 0 & 0 & 0\end{array}\right]$. (Verify.) Thus if $\mathbf{p} = \begin{bmatrix}x_1\\x_2\end{bmatrix}$ we have $x_1 = ix_2$ so $\mathbf{p} = \begin{bmatrix}ix_2\\x_2\end{bmatrix} = x_2\begin{bmatrix}i\\1\end{bmatrix}$, $x_2 \neq 0$. Choosing $x_2 = 1$ gives eigenvector $\mathbf{x} = \begin{bmatrix}i\\1\end{bmatrix}$.

Case $\lambda_2 = \frac{\sqrt{2} - i\sqrt{2}}{2}$: By calculations similar to the previous case we get eigenpair $\mathbf{x} = \begin{bmatrix} -i \\ 1 \end{bmatrix}$. (Verify.)

■

Example 3.

Let

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{bmatrix}.$$

The characteristic polynomial of A is (verify)

$$f(\lambda) = \det(\lambda I_3 - A) = \begin{vmatrix} \lambda - 1 & -2 & 1 \\ -1 & \lambda - 0 & -1 \\ -4 & 4 & \lambda - 5 \end{vmatrix}$$

$$= \lambda^3 - 6\lambda^2 + 11\lambda - 6.$$

Factoring : $\lambda^3 - 6\lambda^2 + 11\lambda - 6$, we have

$$f(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3).$$

The eigenvalues of A are then

$$\lambda_1 = 1, \qquad \lambda_2 = 2, \qquad \lambda_3 = 3.$$

To find an eigenvector $\mathbf{x}_1$ associated with $\lambda_1 = 1$, we form the linear system

$$(1I_3 - A)\mathbf{x} = \mathbf{0},$$

$$\begin{bmatrix} 1-1 & -2 & 1 \\ -1 & 1 & -1 \\ -4 & 4 & 1-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 0 & -2 & 1 \\ -1 & 1 & -1 \\ -4 & 4 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The rref of the augmented matrix is $\left[\begin{array}{ccc|c} 1 & 0 & 1/2 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$.

A solution is

$$\begin{bmatrix} -\frac{1}{2}r \\ \frac{1}{2}r \\ r \end{bmatrix}$$

for any number r. Thus, for $r = 2$,

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

is an eigenvector of A associated with $\lambda_1 = 1$.

To find an eigenvector $\mathbf{x}_2$ associated with $\lambda_2 = 2$, we form the linear system

$$(2I_3 - A)\mathbf{x} = \mathbf{0},$$

that is,

$$\begin{bmatrix} 2-1 & -2 & 1 \\ -1 & 2 & -1 \\ -4 & 4 & 2-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 1 & -2 & 1 \\ -1 & 2 & -1 \\ -4 & 4 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

A solution is

$$\begin{bmatrix} -\frac{1}{2}r \\ \frac{1}{4}r \\ r \end{bmatrix}$$

for any number r. Thus, for $r = 4$,

$$\mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}$$

rref is $\left[\begin{array}{ccc|c} 1 & 0 & 1/2 & 0 \\ 0 & 1 & -1/4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

is an eigenvector of A associated with $\lambda_2 = 2$.

To find an eigenvector $\mathbf{x}_3$ associated with $\lambda_3 = 3$, we form the linear system

$$(3I_3 - A)\mathbf{x} = \mathbf{0},$$

and find that a solution is (verify)

$$\begin{bmatrix} -\frac{1}{4}r \\ \frac{1}{4}r \\ r \end{bmatrix}$$

for any number r. Thus, for $r = 4$,

$$\mathbf{x}_3 = \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix}$$

is an eigenvector of A associated with $\lambda_3 = 3$. ■

SOME FACTS:

An $n \times n$ matrix A is singular if and only if 0 is an eigenvalue of A.

The eigenvalues of A are the roots of the characteristic polynomial of A.

We now extend our list of nonsingular equivalences.

List of Nonsingular Equivalences

The following statements are equivalent for an $n \times n$ matrix A.

1. A is nonsingular.
2. $\mathbf{x} = \mathbf{0}$ is the only solution to $A\mathbf{x} = \mathbf{0}$.
3. A is row equivalent to I_n.
4. The linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $n \times 1$ matrix $\mathbf{b}$.
5. $\det(A) \neq 0$.
6. A has rank n.
7. A has nullity 0.
8. The rows of A form a linearly independent set of n vectors in R^n.
9. The columns of A form a linearly independent set of n vectors in R^n.
10. Zero is *not* an eigenvalue of A.

EIGENSPACE 🡐