Chapter 5 Eigenvalues and Eigenvectors

Section 5.1: Eigenvalues and Eigenvectors

Section 5.2: The Characteristic Equation

Section 5.3: Diagonalization

Eigenvalues and Eigenvectors

- **Definition:** An **eigenvector** of an $n \times n$ matrix *A* is a nonzero vector **x** such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. A scalar λ is called an **eigenvalue** of *A* if there is a nontrivial solution **x** of $A\mathbf{x} = \lambda\mathbf{x}$; such an **x** is called an **eigenvector corresponding to λ.**
- λ is an eigenvalue of an $n \times n$ matrix *A* if and only if the equation

$$(A - \lambda I)\mathbf{x} = 0 \qquad (1)$$

has a nontrivial solution.
- The set of *all* solutions of (1) is just the null space of the matrix $A - \lambda I$

Eigenvalues and Eigenvectors

- So this set is a *subspace* of $\mathbb{R}^n$ and is called the **eigenspace** of *A* corresponding to λ.

- The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ.

- **Example 1:** Show that 7 is an eigenvalue of matrix

$$A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$$ and find the corresponding eigenvectors.

Eigenvalues and Eigenvectors

- **Solution:** The scalar 7 is an eigenvalue of *A* if and only if the equation

$$A\mathbf{x} = 7\mathbf{x} \qquad (2)$$

has a nontrivial solution.

- But (2) is equivalent to $A\mathbf{x} - 7\mathbf{x} = \mathbf{0}$, or

$$(A - 7I)\mathbf{x} = \mathbf{0} \qquad (3)$$

- To solve this homogeneous equation, form the matrix

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

Eigenvalues and Eigenvectors

- The columns of $A-7I$ are obviously linearly dependent, so (3) has nontrivial solutions.
- To find the corresponding eigenvectors, use row operations:

$$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- The general solution has the form $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
- Each vector of this form with $x_2 \neq 0$ is an eigenvector corresponding to $\lambda = 7$.

Eigenvalues and Eigenvectors

- **Example 2:** Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. An eigenvalue of *A* is 2. Find a basis for the corresponding eigenspace.
- **Solution:** Form

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$$

and row reduce the augmented matrix for $(A - 2I)\mathbf{x} = \mathbf{0}$.

Eigenvalues and Eigenvectors

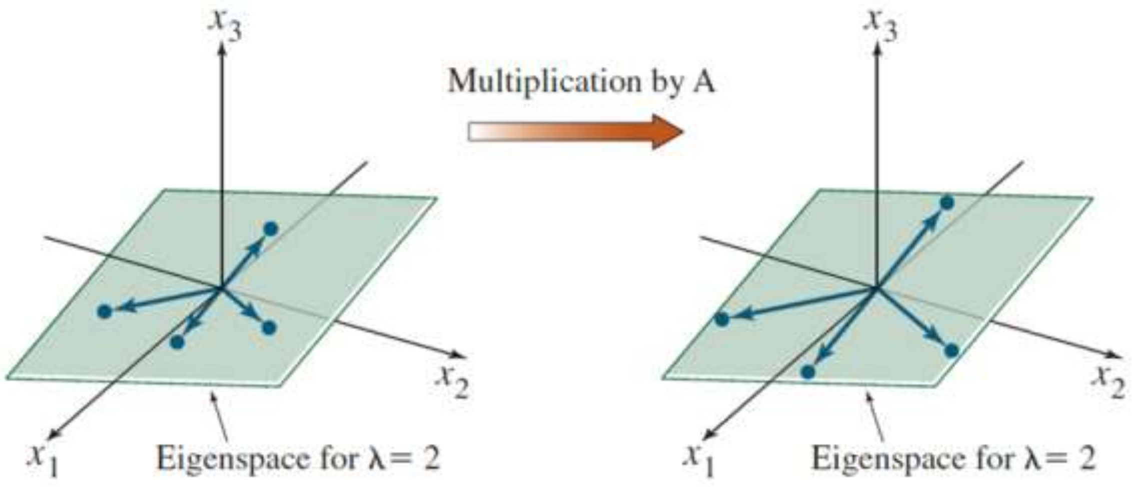
$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- At this point, it is clear that 2 is indeed an eigenvalue of A because the equation $(A-2I)\mathbf{x}=\mathbf{0}$ has free variables.
- The general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \quad x_2 \text{ and } x_3 \text{ free.}$$

Eigenvalues and Eigenvectors

- The eigenspace, shown in the following figure, is a two-dimensional subspace of $\mathbb{R}^3$.



A acts as a dilation on the eigenspace.

- A basis is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$

Eigenvalues and Eigenvectors

- **Theorem:** The eigenvalues of a triangular matrix are the entries on its main diagonal.
- **Proof:** For simplicity, consider the 3×3 case.

If *A* is upper triangular, the $A-\lambda I$ has the form

$$A-\lambda I=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}-\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$=\begin{bmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ 0 & a_{22}-\lambda & a_{23} \\ 0 & 0 & a_{33}-\lambda \end{bmatrix}$$

Eigenvalues and Eigenvectors

- The scalar λ is an eigenvalue of *A* if and only if the equation $(A-\lambda I)\mathbf{x}=\mathbf{0}$ has a nontrivial solution, that is, if and only if the equation has a free variable.

- Because of the zero entries in $A-\lambda I$, it is easy to see that $(A-\lambda I)\mathbf{x}=\mathbf{0}$ has a free variable if and only if at least one of the entries on the diagonal of $A-\lambda I$ is zero.

- This happens if and only if λ equals one of the entries a_{11}, a_{22}, a_{33} in *A*.

Eigenvalues and Eigenvectors

- **Theorem:** If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.
- **Proof:** Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly dependent.

Since $\mathbf{v}_1$ is nonzero, one of the vectors in the set is a linear combination of the preceding vectors.

Let p be the least index such that $\mathbf{v}_{p+1}$ is a linear combination of the preceding (linearly independent) vectors.

Eigenvalues and Eigenvectors

- Then there exist scalars $c_1, \ldots, c_p$ such that

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{v}_{p+1} \tag{4}$$

- Multiplying both sides of (4) by A and using the fact that

$$c_1A\mathbf{v}_1 + \cdots + c_pA\mathbf{v}_p = A\mathbf{v}_{p+1}$$

$$c_1\lambda_1\mathbf{v}_1 + \cdots + c_p\lambda_p\mathbf{v}_p = \lambda_{p+1}\mathbf{v}_{p+1} \tag{5}$$

- Multiplying both sides of (4) by λ_{p+1} and subtracting the result from (5), we have

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \cdots + c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = 0 \tag{6}$$

Eigenvalues and Eigenvectors

- Since {$\mathbf{v}_1, \ldots, \mathbf{v}_p$} is linearly independent, the weights in (6) are all zero.

But none of the factors $\lambda_i - \lambda_{p+1}$ are zero, because the eigenvalues are distinct.

Hence $c_i = 0$ for $i = 1, \ldots, p$.

But then (4) says that $\mathbf{v}_{p+1} = \mathbf{0}$, which is impossible.

Hence {$\mathbf{v}_1, \cdots, \mathbf{v}_r$} cannot be linearly dependent and therefore must be linearly independent.

Section 5.1: Eigenvalues and Eigenvectors

Section 5.2: The Characteristic Equation

Section 5.3: Diagonalization

Review of Determinants

- Let A be a 2×2 matrix, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. The **determinant** of A is given by $\det A = a_{11}a_{22} - a_{12}a_{21}$.
- The matrix A_{ij} is formed from the matrix A by removing the i-th row and j-th column of A. Then

$$C_{ij} = (-1)^{i+j} \det A_{ij}$$

- The expansion across the i th row is

$$\det A = ai_1 C_{i1} + ai_2 C_{i2} + \cdots + a_{in} C_{in}$$

- The cofactor expansion down the j th column is

$$\det A = a_{1j} C_{1j} + a_{2j} C_{2j} + \cdots + anj C_{nj}$$

Review of Determinants

- **Theorem: Properties of Determinants**

  Let *A* and *B* be $n \times n$ matrices.

  a. *A* is invertible if and only if $\det A \neq 0$.

  b. $\det AB = (\det A)(\det B)$.

  c. $\det A^T = \det A$.

  d. If *A* is triangular, then det *A* is the product of the entries on the main diagonal.

The Characteristic Equation

- The scalar equation $\det(A-\lambda I)=0$ is called the **characteristic equation** of *A*.
- A scalar λ is an eigenvalue of an $n \times n$ matrix *A* if and only if λ satisfies the characteristic equation

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

- Note that $\det(A-\lambda I)=0$ if and only if there is nonzero vector **x** in $\operatorname{Nul}(A-\lambda I)$ if and only if there is a nonzero vector **x** such that $A\mathbf{x}=\lambda\mathbf{x}$.

The Characteristic Equation

- **Example:** Find the characteristic equation of

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- **Solution:** Form $A - \lambda I$, and compute

The Characteristic Equation

$$\det(A-\lambda I)=\det\begin{bmatrix}5-\lambda & -2 & 6 & -1\\ 0 & 3-\lambda & -8 & 0\\ 0 & 0 & 5-\lambda & 4\\ 0 & 0 & 0 & 1-\lambda\end{bmatrix}$$

$$=(5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda)$$

- The characteristic equation is

$$(5-\lambda)^2(3-\lambda)(1-\lambda)=0$$

or

$$(\lambda-5)^2(\lambda-3)(\lambda-1)=0$$

The Characteristic Equation

- Expanding the product, we can also write

$$\lambda^4 - 14\lambda^3 + 68\lambda^2 - 130\lambda + 75 = 0$$

- The eigenvalue 5 is said to have **multiplicity** 2 because

$(\lambda - 5)$ occurs two times as a factor of the characteristic polynomial.

- In general, the (algebraic) multiplicity of an eigenvalue λ is its multiplicity as a root of the characteristic equation.

The Invertible Matrix Theorem

- **Theorem:** Let A be an $n \times n$ matrix. Then A is invertible if and only if:

  The number 0 is *not* an eigenvalue of A.

- **Proof:**

A is not invertible

if and only if

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

if and only if

0 is an eigenvalue of A

Similarity

- If *A* and *B* are $n \times n$ matrices, then *A* **is similar to** *B* if there is an invertible matrix *P* such that $P^{-1}AP = B$, or, equivalently,
$$A = PBP^{-1}$$
- Writing *Q* for P^{-1}, we have $Q^{-1}BQ = A$.
- So *B* is also similar to *A*, and we say simply that *A* and *B* are **similar**.
- Changing *A* into $P^{-1}AP$ is called a **similarity transformation.**

Similarity

- **Theorem:** If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

- **Proof:** If $B = P^{-1}AP$, then,

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

- Using the multiplicative property (b) in Theorem (3), we compute

$$\begin{aligned} \det(B - \lambda I) &= \det\left[P^{-1}(A - \lambda I)P\right] \\ &= \det(P^{-1})\det(A - \lambda I)\det(P) \end{aligned} \tag{1}$$

Similarity

- Since $\det(P^{-1})\det(P) = \det(P^{-1}P) = \det I = 1$, we see from equation (1) that $\det(B - \lambda I) = \det(A - \lambda I)$.

- **Warnings:**
  1. The matrices

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \text{ and } \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

     are not similar even though they have the same eigenvalues.

Similarity

2. Similarity is not the same as row equivalence. (If *A* is row equivalent to *B*, then $B = EA$ for some invertible matrix *E*). Row operations on a matrix usually change its eigenvalues.

Section 5.1: Eigenvalues and Eigenvectors

Section 5.2: The Characteristic Equation

Section 5.3: Diagonalization

Diagonalization

- **Example:** Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a formula for

  A^k, given that $A = PDP^{-1}$, where

  $$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \text{ and } D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

- **Solution:**

  $$P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Diagonalization

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD\underbrace{(P^{-1}P)}_{I}DP^{-1} = PDDP^{-1}$$

$$= PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

$$A^3 = (PDP^{-1})A^2 = (PD\underset{I}{P^{-1})P}D^2P^{-1} = PDD^2P^{-1} = PD^3P^{-1}$$

Diagonalization

- In general, for $k \geq 1$,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix}$$

- A square matrix *A* is said to be **diagonalizable** if *A* is similar to a diagonal matrix, that is, if $A = PDP^{-1}$ for some invertible matrix *P* and some diagonal, matrix *D*.

The Diagonalization Theorem

- **Theorem:** An $n \times n$ matrix *A* is diagonalizable if and only if *A* has *n* linearly independent eigenvectors.

  In fact, $A = PDP^{-1}$, with *D* a diagonal matrix, if and only if the columns of *P* are *n* linearly independent eigenvectors of *A*. In this case, the diagonal entries of *D* are eigenvalues of *A* that correspond, respectively, to the eigenvectors in *P*.

  In other words, *A* is diagonalizable if and only if there are enough eigenvectors to form a basis of $\mathbb{R}^n$. We call such a basis an **eigenvector basis** of $\mathbb{R}^n$.

Diagonalizing Matrices

- **Example:** Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

  That is, find an invertible matrix *P* and a diagonal matrix *D* such that $A = PDP^{-1}$.

- **Solution:**
- *Step 1.* Find the eigenvalues of A**.**

Here, the characteristic equation turns out to involve a cubic polynomial that can be factored:

Diagonalizing Matrices

$$0 = \det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4$$
$$= -(\lambda - 1)(\lambda + 2)^2$$

- The eigenvalues are $\lambda = 1$ and $\lambda = -2$.
- *Step 2.* Find three linearly independent eigenvectors of A.
- Three vectors are needed because *A* is a 3×3 matrix.

(This is a critical step. If it fails, then Theorem 5 says that *A* cannot be diagonalized.)

Diagonalizing Matrices

- Basis for $\lambda = 1 : \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$
- Basis for $\lambda = -2 : \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$
- You can check that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set.

Diagonalizing Matrices

- *Step 3.* Construct P from the vectors in step 2.

(The order of the vectors is unimportant.)

- Using the order chosen in step 2, form

$$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

- *Step 4.* Construct D from the corresponding eigenvalues.

(In this step, it is essential that the order of the eigenvalues matches the order chosen for the columns of *P*.)

Diagonalizing Matrices

- Use the eigenvalue $\lambda = -2$ twice, once for each of the eigenvectors corresponding to $\lambda = -2$:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

- To avoid computing P^{-1}, simply verify that $AP = PD$.
- Compute

$$AP = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

Diagonalizing Matrices

$$PD = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

- **Theorem:** An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.
- **Proof:** Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be eigenvectors corresponding to the n distinct eigenvalues of a matrix A.
- Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent.
- Hence A is diagonalizable.

Matrices Whose Eigenvalues are not Distinct

- It is not *necessary* for an $n \times n$ matrix to have *n* distinct eigenvalues in order to be diagonalizable.
- Theorem 6 provides a *sufficient* condition for a matrix to be diagonalizable.
- If an $n \times n$ matrix *A* has *n* distinct eigenvalues, with corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and if $P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}$, then *P* is automatically invertible because its columns are linearly independent, by Theorem 2.

Matrices Whose Eigenvalues are not Distinct

- When A is diagonalizable but has fewer than n distinct eigenvalues, it is still possible to build P in a way that makes P automatically invertible, as the next theorem shows.

- **Theorem:** Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_p$.
  a. For $1 \leq k \leq p$, the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k.

Matrices Whose Eigenvalues are not Distinct

b. The matrix A is diagonalizable if and only if the sum of the dimensions of the eigenspaces equals n, and this happens if and only if (*i*) the characteristic polynomial factors completely into linear factors and (*ii*) the dimension of the eigenspace for each λ_k equals the multiplicity of λ_k.

c. If A is diagonalizable and B_k is a basis for the eigenspace corresponding to λ_k for each k, then the total collection of vectors in the sets $B_1, \ldots, B_p$ forms an eigenvector basis for $\mathbb{R}^n$.