Chapter 5
Eigenvalues and Eigenvectors

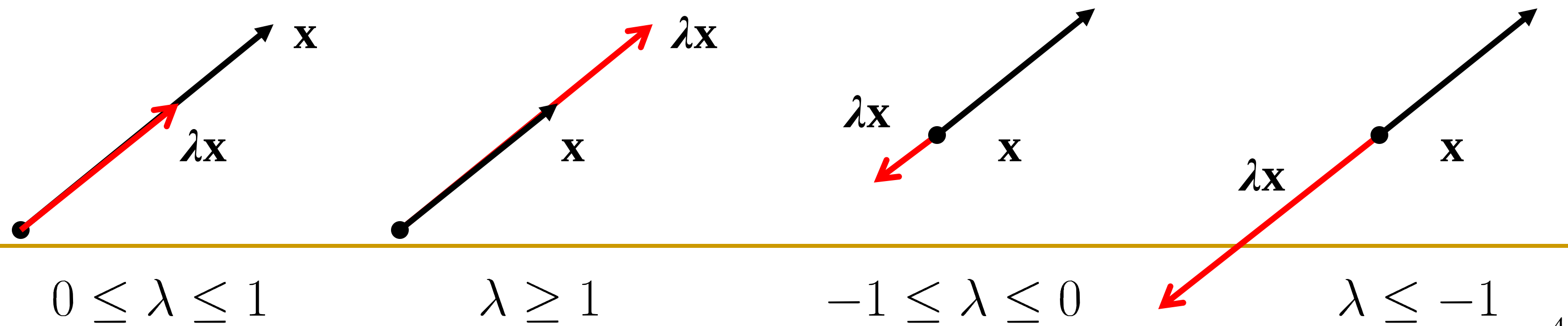
Outline

- 5.1 Eigenvalues and Eigenvectors
- 5.2 Diagonalization
- 5.3 Complex Vector Spaces

5.1
Eigenvalues and Eigenvectors

Eigenvalue and Eigenvector

- If A is a $n \times n$ matrix, then **a nonzero vector x** in R^n is called an **eigenvector** of A (or the matrix operator T_A) if $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$; that is $A\mathbf{x}= \lambda\mathbf{x}$ for some scalar λ. The scalar λ is called an **eigenvalue** of A (or of T_A), and x is said to be an **eigenvector corresponding to λ**.
- In some special case, multiplication by A leaves the direction unchanged.



$0 \le \lambda \le 1$ $\lambda \ge 1$ $-1 \le \lambda \le 0$ $\lambda \le -1$

Example

- The vector $\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$

$$A\boldsymbol{x} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\boldsymbol{x} \qquad \lambda = 3$$

Computing Eigenvalues and Eigenvectors

- To compute eigenvalues and eigenvectors

$$A\boldsymbol{x} = \lambda \boldsymbol{x} \Rightarrow A\boldsymbol{x} = \lambda I \boldsymbol{x} \Rightarrow (\lambda I - A)\boldsymbol{x} = 0$$

- For λ to be an eigenvalue of A **this equation must have a nonzero solution** for **x**. from Theorem 4.9.4, this is so if and only if the coefficient matrix $\lambda I - A$ has a zero determinant.
- Theorem 5.1.1: If A is an $n \times n$ matrix, then λ is an eigenvalue of A if and only if it satisfies the equation:

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

- This is called the **characteristic equation** of A.

Example

- Find all eigenvalues of the matrix $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$
- Characteristic function:

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{vmatrix} = 0$$

$$(\lambda - 3)(\lambda + 1) = 0 \qquad \lambda = 3 \qquad \lambda = -1$$

Characteristic Polynomial

- When the determinant det(λI-A) is expanded, the result is a polynomial $p(\lambda)$ of degree n that is called the **characteristic polynomial** of A.
- Example:

$$p(\lambda) = (\lambda - 3)(\lambda + 1) = \lambda^2 - 2\lambda - 3$$

- The characteristic polynomial of an $n \times n$ matrix

$$p(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0$$

- Has **at most** n distinct eigenvalues.
- Some solutions may be complex numbers; it is possible for a matrix to have complex eigenvalues.

Example

- Find eigenvalues of $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$

$$\det(\lambda I - A) = \begin{vmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -4 & 17 & \lambda - 8 \end{vmatrix} = \lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0$$

$$(\lambda - 4)(\lambda^2 - 4\lambda + 1) = 0$$

$$\lambda = 4 \qquad \lambda = 2 + \sqrt{3} \qquad \lambda = 2 - \sqrt{3}$$

Example

- Find eigenvalues of $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} & -a_{14} \\ 0 & \lambda - a_{22} & -a_{23} & -a_{24} \\ 0 & 0 & \lambda - a_{33} & -a_{34} \\ 0 & 0 & 0 & \lambda - a_{14} \end{vmatrix}$$

$$= (\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44}) = 0$$

$$\lambda = a_{11} \qquad \lambda = a_{22} \qquad \lambda = a_{33} \qquad \lambda = a_{44}$$

- Theorem 5.1.2: If A is an $n \times n$ triangular matrix (upper triangular, lower triangular, or diagonal), then the eigenvalues of A are the entries on the main diagonal of A.

Theorem 5.1.3

- If A is an $n \times n$ matrix, the following statements are equivalent
  - λ is an eigenvalue of A
  - The system of equations $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has nontrivial solutions
  - There is a nonzero vector x such that $A\mathbf{x} = \lambda\mathbf{x}$
  - λ is a solution of the characteristic equation $\det(\lambda I - A) = 0$

Finding Eigenvectors and Bases for Eigenspaces

- The eigenvectors are the nonzero vectors in the null space of the matrix $\lambda I\text{-}A$

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

- We call this null space the eigenspace of A corresponding to λ.
- The eigenspace of A corresponding to the eigenvalue λ is the solution space of the homogeneous system $(\lambda I\text{-}A)\mathbf{x}=\mathbf{0}$.

Example

- Eigenvalues of the matrix $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$ are $\begin{matrix} \lambda = 3 \\ \lambda = -1 \end{matrix}$
- The system (λI-A)**x**=**0**

$$\begin{bmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

⇨ $\lambda = 3$ $$\begin{bmatrix} 0 & 0 \\ -8 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$ $$\begin{aligned} x_1 &= \tfrac{1}{2}t \\ x_2 &= t \end{aligned}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}t \\ t \end{bmatrix} = t \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$$

$\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$ is a basis for the eigenspace corresponding to $\lambda = 3$

Example

- Find bases for the eigenspaces of $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$
- Solution:
  - The characteristic equation of matrix A is $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$, or in factored form, $(\lambda - 1)(\lambda - 2)^2 = 0$; thus, the eigenvalues of A are $\lambda = 1$ and $\lambda = 2$, so there are two eigenspaces of A.
  - $(\lambda I - A)\mathbf{x} = \mathbf{0} \Rightarrow$ $$\begin{bmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (3)$$
  - If $\lambda = 2$, then (3) becomes $$\begin{bmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Example

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

- Solving the system yield

$$x_1 = -s,\ x_2 = t,\ x_3 = s$$

- Thus, the eigenvectors of A corresponding to $\lambda = 2$ are the nonzero vectors of the form

$$\mathbf{x} = \begin{bmatrix} -s \\ t \\ s \end{bmatrix} = \begin{bmatrix} -s \\ 0 \\ s \end{bmatrix} + \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} = s \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

- The vectors $[-1\ 0\ 1]^T$ and $[0\ 1\ 0]^T$ are linearly independent and form a basis for the eigenspace corresponding to $\lambda = 2$.
- Similarly, the eigenvectors of A corresponding to $\lambda = 1$ are the nonzero vectors of the form $\mathbf{x} = s\,[-2\ 1\ 1]^T$
- Thus, $[-2\ 1\ 1]^T$ is a basis for the eigenspace corresponding to $\lambda = 1$.

Powers of a Matrix

- If λ is an eigenvalue of A and $\mathbf{x}$ is a corresponding eigenvector, then

$$A^2\boldsymbol{x} = A(A\boldsymbol{x}) = A(\lambda\boldsymbol{x}) = \lambda(A\boldsymbol{x}) = \lambda(\lambda\boldsymbol{x}) = \lambda^2\boldsymbol{x}$$

  which shows that λ^2 is an eigenvalue of A^2 and that $\mathbf{x}$ is a corresponding eigenvector.

Theorems

- Theorem 5.1.4
  - If k is a positive integer, λ is an eigenvalue of a matrix A, and $\mathbf{x}$ is corresponding eigenvector, then λ^k is an eigenvalue of A^k and $\mathbf{x}$ is a corresponding eigenvector.

- Theorem 5.1.5
  - A square matrix A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.

Proof of Theorem 5.1.5

- Assume that A is an $n \times n$ matrix and observe that λ=0 is a solution of the characteristic equation
$$\lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0$$
if and only if the constant term c_n is zero.
- Thus it suffices to prove that A is invertible if and only if $c_n \neq 0$. But $\det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0$
or, on setting λ=0,
$$\det(-A) = c_n \quad \text{or} \quad (-1)^n \det(A) = c_n$$
- It follows from the last equation that det(A)=0 if and only if c_n=0, and this in turn implies that A is invertible if and only if $c_n \neq 0$.

Theorem 5.1.6 (Equivalent Statements)

- If A is an $m \times n$ matrix, and if $T_A : R^n \rightarrow R^n$ is multiplication by A, then the following are equivalent:
  - A is invertible.
  - $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  - The reduced row-echelon form of A is I_n.
  - A is expressible as a product of elementary matrices.
  - $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.
  - $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.
  - $\det(A) \neq 0$.
  - The column vectors of A are linearly independent.
  - The row vectors of A are linearly independent.
  - The column vectors of A span R^n.
  - The row vectors of A span R^n.

Theorem 5.1.6 (Equivalent Statements)

- The column vectors of A form a basis for R^n.
- The row vectors of A form a basis for R^n.
- A has rank n.
- A has nullity 0.
- The orthogonal complement of the nullspace of A is R^n.
- The orthogonal complement of the row space of A is $\{\mathbf{0}\}$.
- The range of T_A is R^n.
- T_A is one-to-one.
- A^TA is invertible.
- $\lambda = 0$ is not an eigenvalue of A.

$A\mathbf{x} = \lambda\mathbf{x}$
=> $(\lambda I - A)\mathbf{x} = \mathbf{0}$
=> $-A\mathbf{x} = 0$ (if $\lambda = 0$)
$\Rightarrow A\mathbf{x} = 0$
$\Rightarrow$ **x** has the only trivial solution i.e., $\mathbf{x} = \mathbf{0}$
$\Rightarrow$ (since A is invertable)

5.2 Diagonalization

Matrix Diagonalization Problem

- Problem 1: Given an $n \times n$ matrix A, does there exist an invertible matrix P such that $P^{-1}AP$ is diagonal?
- Problem 2: Given an $n \times n$ matrix A, does A have n linearly independent eigenvectors?

- The matrix product $P^{-1}AP$ in Problem 1 is called a ***similarity transformation*** of the matrix A.
- If A and B are square matrices, then we say that **B is similar to A** if there is an invertible matrix P such that $B = P^{-1}AP$

Similarity Invariants

- If B is similar to A, then A is similar to B, since we can express B as $B=Q^{-1}AQ$ by taking $Q=P^{-1}$.
- We usually say that A and B are similar matrices.
- If $B=P^{-1}AP$, then A and B have the same determinant

$$\det(B)=\det(P^{-1}AP)=\det(P^{-1})\det(A)\det(P) = (1/\det(P))\det(A)\det(P) = \det(A)$$

- Any property that is shared by all similar matrices is called a **similarity invariant** or is said to be **invariant under similarity**.

Similarity Invariants

Property	Description
Determinant	A and $P^{-1}AP$ have the same determinant.
Invertibility	A is invertible if and only if $P^{-1}AP$ is invertible.
Rank	A and $P^{-1}AP$ have the same rank.
Nullity	A and $P^{-1}AP$ have the same nullity.
Trace	A and $P^{-1}AP$ have the same trace.
Characteristic polynomial	A and $P^{-1}AP$ have the same characteristic polynomial.
Eigenvalues	A and $P^{-1}AP$ have the same eigenvalues.
Eigenspace dimension	If λ is an eigenvalue of A and hence of $P^{-1}AP$, then the eigenspace of A corresponding to λ and the eigenspace of $P^{-1}AP$ corresponding to λ **have the same dimension**. (have the same set of eigenvectors?)

Diagonalizable

- A square matrix A is said to be ***diagonalizable*** if it is similar to some diagonal matrix; that is, if there exists an invertible matrix P such that $P^{-1}AP$ is diagonal. In this case the matrix P is said to ***diagonalize*** (對角化) A.
- Theorem 5.2.1: If A is an $n \times n$ matrix, the following statements are equivalent.
  - (a) A is diagonalizable
  - (b) A has n linearly independent eigenvectors.

Proof of Theorem 5.2.1

- Since A is assumed diagonalizable, there is an invertible matrix

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \qquad D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

  such that $P^{-1}AP$ is diagonal, say $P^{-1}AP = D$.

- It follows that $AP = PD$; that is

$$AP = PD = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \cdots & \lambda_n p_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \cdots & \lambda_n p_{nn} \end{bmatrix}$$

Proof of Theorem 5.2.1

$$AP = PD = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \cdots & \lambda_n p_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \cdots & \lambda_n p_{nn} \end{bmatrix}$$

- If we now let $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ denote the column vectors of P, then the successive columns of AP are $\lambda_1 \boldsymbol{p}_1, \lambda_2 \boldsymbol{p}_2, \cdots, \lambda_n \boldsymbol{p}_n$
- We also know that the successive columns of AP are $A\mathbf{p}_1, A\mathbf{p}_2, \ldots, A\mathbf{p}_n$. Thus we have

$$A\boldsymbol{p}_1 = \lambda_1 \boldsymbol{p}_1, A\boldsymbol{p}_2 = \lambda_2 \boldsymbol{p}_2, \cdots, A\boldsymbol{p}_n = \lambda_n \boldsymbol{p}_n$$

- Since P is invertible, its columns are all nonzero. Thus, $\lambda_1, \lambda_2, \cdots, \lambda_n$ are eigenvalues of A, and $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are corresponding eigenvectors.
- Since P is invertible, $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are linearly independent. Thus A has n linearly independent eigenvectors.

Procedure for Diagonalizing a Matrix

- The preceding theorem guarantees that an $n \times n$ matrix A with n linearly independent eigenvectors is diagonalizable, and the proof provides the following method for diagonalizing A.
  - **Step 1.** Find n linear independent eigenvectors of A, say, $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$.
  - **Step 2.** From the matrix P having $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ as its column vectors.
  - **Step 3.** The matrix $P^{-1}AP$ will then be diagonal with $\lambda_1, \lambda_2, \ldots, \lambda_n$ as its successive diagonal entries, where λ_i is the eigenvalue corresponding to $\mathbf{p}_i$, for $i = 1, 2, \ldots, n$.
- If there is total of n such vectors, then A is diagonalizable. Otherwise, A is not diagonalizable.

Example

- Find a matrix P that diagonalizes $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$
- Solution:
  - From the previous example, we have the following bases for the eigenspaces:
    - $\lambda = 2: \quad \mathbf{p}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \lambda = 1: \quad \mathbf{p}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$
  - Thus,

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

  - Also,

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

Example

- There is no preferred order for the columns of P. Since the ith diagonal entry of $P^{-1}AP$ is an eigenvalue for the ith column vector of P, changing the order of the columns of P just changes the order of the eigenvalues of the diagonal of $P^{-1}AP$.
- If we write P as

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

We have

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Example (A Non-Diagonalizable Matrix)

- Find a matrix P that diagonalizes $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{bmatrix}$
- Solution:
  - The characteristic polynomial of A is

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & 0 & 0 \\ -1 & \lambda - 2 & 0 \\ 3 & -5 & \lambda - 2 \end{vmatrix} = (\lambda - 1)(\lambda - 2)^2$$

  - The bases for the eigenspaces are
    - $\lambda = 1$: $\mathbf{p}_1 = \begin{bmatrix} 1/8 \\ -1/8 \\ 1 \end{bmatrix}$ $\lambda = 2$: $\mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$
  - Since there are only two basis vectors in total, A is not diagonalizable.

Example (Alternative Solution)

- If one is interested only in determining whether a matrix is diagonalizable and is not concerned with actually finding a diagonalizing matrix *P*, then it's not necessary to compute bases for the eigenspaces; it suffices to find the dimensions of the eigenspace.
- For this example, the eigenspace corresponding to λ=1 is the solution space of the system
$$\begin{bmatrix} 0 & 0 & 0 \\ -1 & -1 & 0 \\ 3 & -5 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
- The coefficient matrix has rank 2. Thus the nullity of this matrix is 1, and hence the solution space is one-dimensional.

Example (Alternative Solution)

- The eigenspace corresponding to λ=2 is the solution space of the system $\begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 3 & -5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$
- The coefficient matrix also has rank 2 and nullity 1, so the eigenspace corresponding to λ=2 is also one-dimensional.
- Since the eigenspaces produce a total of two basis vectors, the matrix *A* is not diagonalizable.

Theorems

- Theorem 5.2.2
  - If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a linearly independent set.

- Theorem 5.2.3
  - If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Example

- Since the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$$

  has three distinct eigenvalues, $\lambda = 4,\ \lambda = 2+\sqrt{3},\ \lambda = 2-\sqrt{3}$
- Therefore, *A* is diagonalizable.
- Further,

$$P^{-1}AP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2+\sqrt{3} & 0 \\ 0 & 0 & 2-\sqrt{3} \end{bmatrix}$$

  for some invertible matrix *P*, and the matrix *P* can be found using the procedure for diagonalizing a matrix.

A Diagonalizable Matrix

- Since the eigenvalues of a triangular matrix are the entries on its main diagonal (Theorem 5.1.2).
- Thus, a triangular matrix with distinct entries on the main diagonal is diagonalizable.

- For example,

$$A = \begin{bmatrix} -1 & 2 & 4 & 0 \\ 0 & 3 & 1 & 7 \\ 0 & 0 & 5 & 8 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

is a diagonalizable matrix.

Computing Powers of a Matrix

- If A is an $n \times n$ matrix and P is an invertible matrix, then

$$(P^{-1}AP)^2 = P^{-1}APP^{-1}AP = P^{-1}AIAP = P^{-1}A^2P$$

- $(P^{-1}AP)^k = P^{-1}A^kP$ for any positive integer k.
- If A is diagonalizable, and $P^{-1}AP = D$ is a diagonal matrix, then

$$P^{-1}A^kP = (P^{-1}AP)^k = D^k$$

- Thus,

$$A^k = PD^kP^{-1}$$

- The matrix D^k is easy to compute; for example, if

$$D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix}, \text{ and } \mathrm{D}^{\mathrm{k}} = \begin{bmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & d_n^k \end{bmatrix}$$

Example

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

- Find A^{13}
- The matrix A is diagonalized by $P = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$

  and that $D = P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- Thus,

$$A^{13} = PD^{13}P^{-1} = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2^{13} & 0 & 0 \\ 0 & 2^{13} & 0 \\ 0 & 0 & 1^{13} \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} -8190 & 0 & -16382 \\ 8191 & 8192 & 8191 \\ 8191 & 0 & 16383 \end{bmatrix}$$

Theorem 5.2.4

- If λ is an eigenvalue of a square matrix A and $\mathbf{x}$ is a corresponding eigenvector, and if k any positive integer, then λ^k is an eigenvalue of A^k and $\mathbf{x}$ is a corresponding eigenvector.

- Example:

$$A^2\mathbf{x} = A(A\mathbf{x})=A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}$$

Repeated Eigenvalues and Diagonalizability

- If a matrix has all distinct eigenvalues, then it is diagonalizable. Matrices with repeated eigenvalues *might* be nondiagonalizable.
- For example,

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

  has repeated eigenvalues (1) but is diagonalizable since any nonzero vector in R^3 is an eigenvector of I_3, and so, we can find three linearly independent eigenvectors.

Geometric and Algebraic Multiplicity

- Example: the characteristic polynomial $(\lambda\text{-}1)(\lambda\text{-}2)^2$
  - The eigenspace corresponding to λ=1 is at most one-dim, and the eigenspace corresponding to λ=2 is at most two-dim.
- Definition
  - If λ_0 is an eigenvalue of an $n \times n$ matrix A, then the dimension of the eigenspace corresponding to λ_0 is called the geometric multiplicity (幾何重數) of λ_0, and the number of times that $\lambda - \lambda_0$ appears as a factor in the characteristic polynomial of A is called the algebraic multiplicity (代數重數) of A.

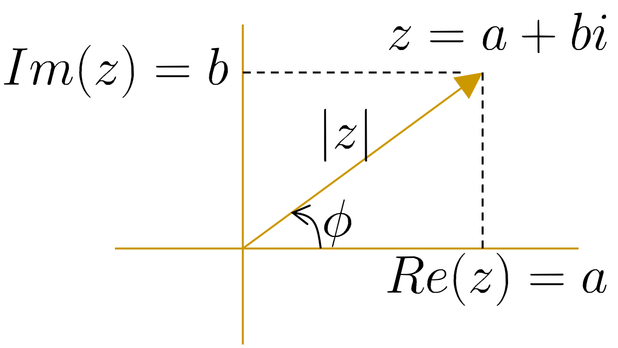
Theorem 5.2.5

- Theorem 5.2.5 (Geometric and Algebraic Multiplicity)
  - If *A* is a square matrix, then :
    - For every eigenvalue of *A* the geometric multiplicity is less than or equal to the algebraic multiplicity.
    - *A* is diagonalizable if and only if the geometric multiplicity is equal to the algebraic multiplicity for every eigenvalue.

5.3
Complex Vector Spaces

Review of Complex Numbers

- If $z = a + bi$ is a complex number, then
  - $Re(z) = a$ and $Im(z) = b$ are called the **real part** of z and the **imaginary part** of z, respectively.
  - $|z| = \sqrt{a^2 + b^2}$ is called the **modulus** (or **absolute value**) of z.
  - $\bar{z} = a - bi$ is called the **complex conjugate** of z.
  - $z\bar{z} = a^2 + b^2 = |z|^2$
  - the angle ϕ is called an **argument** of z.
  - $Re(z) = |z| \cos \phi$
  - $Im(z) = |z| \sin \phi$
  - $z = |z|(\cos \phi + i \sin \phi)$ is called the **polar form** of z.



Complex Eigenvalues

- The characteristic equation of a general $n \times n$ matrix A has the form $\lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0$
- It is possible for the characteristic equation to have imaginary solutions
- $A = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$ has characteristic equation

  $$\begin{vmatrix} \lambda + 2 & 1 \\ -5 & \lambda - 2 \end{vmatrix} = \lambda^2 + 1 = 0 \,,$$

  which has imaginary solutions $\lambda = i$ and $\lambda = -i$

Vectors in C^n

- A vector space in which scalars are allowed to be complex numbers is called a **complex vector space**.
- Definition: If *n* is a positive integer, then a complex ***n*-tuple** is a sequence of *n* complex numbers $(v_1, v_2, \ldots, v_n)$. The set of all complex n-tuple is called ***complex n-space*** and is denoted by C^n. Scalars are complex numbers, and the operations of addition, subtraction, and scalar multiplication are performed componentwise.

Vectors in C^n

- If $v_1, v_2, \ldots, v_n$ are complex numbers, then we call **v**=($v_1, v_2, \ldots, v_n$) a **vector** in C^n and $v_1, v_2, \ldots, v_n$ its **components**. Some examples:

$$\boldsymbol{u} = (1+i, -4i, 3+2i) \quad \boldsymbol{v} = (0, i, 5) \quad \boldsymbol{w} = (6-\sqrt{2}i, 9+\tfrac{1}{2}i, \pi i)$$

- Every vector in C^n can be split into real and imaginary parts

$$\boldsymbol{v} = (v_1, v_2, \ldots, v_n) = (a_1 + b_1 i, a_2 + b_2 i, \ldots, a_n + b_n i)$$

$$\boldsymbol{v} = (a_1, a_2, \ldots, a_n) + i(b_1, b_2, \ldots, b_n)$$

$$\boldsymbol{v} = Re(\boldsymbol{v}) + iIm(\boldsymbol{v})$$

Complex conjugate:

$$\bar{\boldsymbol{v}} = (\bar{v_1}, \bar{v_2}, \ldots, \bar{v_n}) = (a_1 - b_1 i, a_2 - b_2 i, \ldots, a_n - b_n i)$$

$$\boldsymbol{v} = Re(\boldsymbol{v}) - iIm(\boldsymbol{v})$$

Vectors in C^n

- The vectors in R^n can be viewed as those vectors in C^n whose imaginary part is zero. A vector **v** in C^n is in R^n if and only if $\bar{\boldsymbol{v}} = \boldsymbol{v}$
- Real matrix: entries in the matrix are required to be real numbers
- Complex matrix: entries in the matrix are allowed to be complex numbers
- If A is a complex matrix, then Re(A) and Im(A) are the matrices formed from the real and imaginary parts of the entries of A.

Example

- Let v=(3+*i*, -2*i*, 5) $A = \begin{bmatrix} 1+i & -i \\ 4 & 6-2i \end{bmatrix}$
- Then

$$\bar{\boldsymbol{v}} = (3-i, 2i, 5) \qquad Re(\boldsymbol{v}) = (3,0,5) \quad Im(\boldsymbol{v}) = (1,-2,0)$$

$$\bar{A} = \begin{bmatrix} 1-i & i \\ 4 & 6+2i \end{bmatrix} \qquad Re(A) = \begin{bmatrix} 1 & 0 \\ 4 & 6 \end{bmatrix} \qquad Im(A) = \begin{bmatrix} 1 & -1 \\ 0 & -2 \end{bmatrix}$$

$$\det(A) = \begin{vmatrix} 1+i & -i \\ 4 & 6-2i \end{vmatrix} = (1+i)(6-2i) - (-i)(4) = 8+8i$$

Algebraic Properties of the Complex Conjugate

- Theorem 5.3.1: If **u** and **v** are vectors in C^n, and if k is a scalar, then
  - (a) $\bar{\bar{\boldsymbol{u}}} = \boldsymbol{u}$
  - (b) $\overline{k\boldsymbol{u}} = \bar{k}\bar{\boldsymbol{u}}$
  - (c) $\overline{\boldsymbol{u}+\boldsymbol{v}} = \bar{\boldsymbol{u}} + \bar{\boldsymbol{v}}$
  - (d) $\overline{\boldsymbol{u}-\boldsymbol{v}} = \bar{\boldsymbol{u}} - \bar{\boldsymbol{v}}$
- Theorem 5.3.2: If A is an $m \times k$ complex matrix and B is a $k \times n$ complex matrix, then
  - (a) $\bar{\bar{A}} = A$
  - (b) $\overline{A^T} = (\bar{A})^T$
  - (c) $\overline{AB} = \bar{A}\bar{B}$

The Complex Euclidean Inner Product

- Definition: If **u**=(u_1, u_2, …, u_n) and **v**=(v_1, v_2, …, v_n) are vectors in C^n, then the **complex Euclidean inner product** of **u** and **v** (also called **the complex dot product**) id denoted by **u** · **v** and is defined as

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1\bar{v}_1 + u_2\bar{v}_2 + \cdots + u_n\bar{v}_n$$

- We also define the **Euclidean norm** on C^n to be

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}} = \sqrt{|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2}$$

- We call **v** a unit vector in C^n if ||**v**||=1, and we say two vectors **u** and **v** are orthogonal if **u** · **v** =0

Example

- Find **u** · **v**, **v** · **u**, ||**u**||, and ||**v**|| for the vectors

$$\boldsymbol{u} = (1+i, i, 3-i) \qquad \boldsymbol{v} = (1+i, 2, 4i)$$

$$\boldsymbol{u} \cdot \boldsymbol{v} = (1+i)(1-i) + i2 + (3-i)(-4i) = -2 - 10i$$

$$\boldsymbol{v} \cdot \boldsymbol{u} = (1+i)(1-i) + 2(-i) + (4i)(3+i) = -2 + 10i$$

$$\|\boldsymbol{u}\| = \sqrt{|1+i|^2 + |i|^2 + |3-i|^2} = \sqrt{2+1+10} = \sqrt{13}$$

$$\|\boldsymbol{v}\| = \sqrt{|1+i|^2 + |2|^2 + |4i|^2} = \sqrt{2+4+16} = \sqrt{22}$$

Theorem 5.3.3

- In R^n, $\boldsymbol{u}\cdot\boldsymbol{v}=\boldsymbol{u}^T\boldsymbol{v}=\boldsymbol{v}^T\boldsymbol{u}$
- In C^n, $\boldsymbol{u}\cdot\boldsymbol{v}=\boldsymbol{u}^T\bar{\boldsymbol{v}}=\bar{\boldsymbol{v}}^T\boldsymbol{u}$
- Theorem 5.3.3: If **u**, **v**, and **w** are vectors in C^n, and if k is a scalar, then the complex Euclidean inner product has the following properties:
  - (a) $\boldsymbol{u}\cdot\boldsymbol{v}=\overline{\boldsymbol{v}\cdot\boldsymbol{u}}$ **[Antisymmetry property]**
  - (b) $\boldsymbol{u}\cdot(\boldsymbol{v}+\boldsymbol{w})=\boldsymbol{u}\cdot\boldsymbol{v}+\boldsymbol{u}\cdot\boldsymbol{w}$ **[Distributive property]**
  - (c) $k(\boldsymbol{u}\cdot\boldsymbol{v})=(k\boldsymbol{u})\cdot\boldsymbol{v}$ **[Homogeneity property]**
  - (d) $\boldsymbol{u}\cdot k\boldsymbol{v}=\bar{k}(\boldsymbol{u}\cdot\boldsymbol{v})$ **[Antihomogeneity property]**
  - (e) $\boldsymbol{v}\cdot\boldsymbol{v}\geq 0$ and $\boldsymbol{v}\cdot\boldsymbol{v}=0$ if and only if $\boldsymbol{v}=\mathbf{0}$ **[Positivityproperty]**

Proof of Theorem 5.3.3(d)

- $k(\boldsymbol{u}\cdot\boldsymbol{v}) = k(\overline{\boldsymbol{v}\cdot\boldsymbol{u}}) = \bar{\bar{k}}(\overline{\boldsymbol{v}\cdot\boldsymbol{u}})$
  $= \overline{\bar{k}(\boldsymbol{v}\cdot\boldsymbol{u})} = \overline{(\bar{k}\boldsymbol{v})\cdot\boldsymbol{u}} = \boldsymbol{u}\cdot(\bar{k}\boldsymbol{v})$
- To complete this proof, substitute $\bar{k}$ for *k* and use the factor that $\bar{\bar{k}} = k$

Vector Concepts in C^n

- Except for the use of complex scalars, the notions of linear combination, linear independence, subspace, spanning, basis, and dimension carry over without change to C^n.
- If A is an $n \times n$ matrix with complex entries, then the complex roots of the characteristic equation $\det(\lambda I - A) = 0$ are called ***complex eigenvalues*** of A.
- λ is a complex eigenvalue of A iff there exists a nonzero vector **x** in C^n such that $A\boldsymbol{x} = \lambda\boldsymbol{x}$. **x** is called a ***complex eigenvector*** of A corresponding to λ .

Vector Concepts in C^n

- The complex eigenvectors of A corresponding to λ are the nonzero solutions of the linear system $(\lambda I - A)\boldsymbol{x} = \mathbf{0}$ and the set of all such solutions is a subspace of C^n, called the ***eigenspaces*** of A corresponding to λ.

Theorem 5.3.4

- If λ is an eigenvalue of a real $n \times n$ matrix A, and if **x** is a corresponding eigenvector, then $\bar{\lambda}$ is also an eigenvalue of A, and $\bar{\boldsymbol{x}}$ is a corresponding eigenvector.
- Proof:
  - Since λ is a eigenvalue of A and **x** is a corresponding eigenvector, we have $\overline{A\boldsymbol{x}} = \overline{\lambda \boldsymbol{x}} = \bar{\lambda}\bar{\boldsymbol{x}}$
  - However, $\bar{A} = A$, it follows from part (c) of Theorem 5.3.2

$$\overline{A\boldsymbol{x}} = \bar{A}\bar{\boldsymbol{x}} = A\bar{\boldsymbol{x}}$$

  - Therefore, $A\bar{\boldsymbol{x}} = \overline{A\boldsymbol{x}} = \bar{\lambda}\bar{\boldsymbol{x}}$, in which $\bar{\boldsymbol{x}} \neq \mathbf{0}$

Example

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

- Find the eigenvalues and bases for the eigenspace
- Solution:

$$\begin{vmatrix} \lambda + 2 & 1 \\ -5 & \lambda - 2 \end{vmatrix} = \lambda^2 + 1 = (\lambda - i)(\lambda + i)$$

  - Eigenvalues: $\lambda = i$ and $\lambda = -i$
  - To find eigenvectors, we must solve the system

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

  - Solve this system by reducing the augmented matrix $\begin{bmatrix} i + 2 & 1 & 0 \\ -5 & i - 2 & 0 \end{bmatrix}$ by Gauss-Jordan elimination

Example

- The reduced row echelon form must have a row of zeros because it has nontrivial solutions.
- Each row must be a scalar multiple of the other, and hence the first row can be made into a row of zeros by adding a suitable multiple of the second row to it.
- Accordingly, we can simply set the entries in the first row to zero, then interchange the rows, and then multiply the new first row by $-\frac{1}{5}$ to obtain the reduced row echelon form $\begin{bmatrix} 1 & \frac{2}{5}-\frac{1}{5}i & 0 \\ 0 & 0 & 0 \end{bmatrix}$
- A general solution of the system is $\begin{aligned} x_1 &= (-\tfrac{2}{5}+\tfrac{1}{5}i)t \\ x_2 &= t \end{aligned}$

Example

- This tells us that the eigenspace corresponding to $\lambda = i$ is one-dimensional and consists of all complex scalar multiples of the basis vector $\boldsymbol{x} = \begin{bmatrix} -\frac{2}{5} + \frac{1}{5}i \\ 1 \end{bmatrix}$
- As a check, let us confirm $A\mathbf{x}=i\mathbf{x}$.

$$A\boldsymbol{x} = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} -\frac{2}{5} + \frac{1}{5}i \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{5} - \frac{2}{5}i \\ i \end{bmatrix} = i\boldsymbol{x}$$

Example

- We could find the eigenspace corresponding to $\lambda = -i$ in a similar way.

$$\bar{\boldsymbol{x}} = \begin{bmatrix} -\frac{2}{5} - \frac{1}{5}i \\ 1 \end{bmatrix}$$

$$A\bar{\boldsymbol{x}} = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} -\frac{2}{5} - \frac{1}{5}i \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{5} + \frac{2}{5}i \\ -i \end{bmatrix} = -i\bar{\boldsymbol{x}}$$

Eigenvalues and Eigenvectors

- For the 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$
- The characteristic polynomial

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a & -b \\ -c & \lambda - d \end{vmatrix}$$

$$= (\lambda - a)(\lambda - d) - bc = \lambda^2 - (a + d)\lambda + (ad - bc)$$

- We express it as

$$\det(\lambda I - A) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A)$$

Eigenvalues and Eigenvectors

$$\det(\lambda I - A) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A)$$

- Recall that if $ax^2+bx+c=0$ is a quadratic equation with real coefficients, then the discriminant b^2-4ac determines the nature of the roots:
  - $b^2 - 4ac > 0$ [Two distinct real roots]
  - $b^2 - 4ac = 0$ [One repeated real root]
  - $b^2 - 4ac < 0$ [Two conjugate imaginary roots]
- $a = 1$, $b = -\mathrm{tr}(A)$, $c = \det(A)$

Theorem 5.3.5

- If A is a 2×2 matrix with real entries, then the characteristic equation of A is $\lambda^2 - \mathrm{tr}(A)\lambda + \det(A)$ and
  - (a) A has two distinct real eigenvalues if $\mathrm{tr}(A)^2 - 4\det(A) > 0$
  - (b) A has one repeated real eigenvalue if $\mathrm{tr}(A)^2 - 4\det(A) = 0$
  - (c) A has two complex conjugate eigenvalues if $\mathrm{tr}(A)^2 - 4\det(A) < 0$

Example

- Find the eigenvalues of

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

- Solution (a): tr(*A*) = 7 and det(*A*) = 12, so the characteristic equation of *A* is $\lambda^2 - 7\lambda + 12 = 0$

$$(\lambda - 4)(\lambda - 3) = 0 \qquad \lambda = 4 \qquad \lambda = 3$$

- Solution (b): tr(*A*) = 2 and det(*A*) = 1, so the characteristic equation of *A* is $\lambda^2 - 2\lambda + 1 = 0$ $\quad \lambda = 1$
- Solution (c): tr(*A*) = 4 and det(*A*) = 13, so the characteristic equation of *A* is $\lambda^2 - 4\lambda + 13 = 0$

$$\lambda = 2 + 3i$$
$$\lambda = 2 - 3i$$

Symmetric Matrices Have Real Eigenvalues

- Theorem 5.3.6: If *A* is a real symmetric matrix, then *A* has real eigenvalues.
- Proof: Suppose that λ is an eigenvalue of *A* and **x** is a corresponding eigenvector, $A\boldsymbol{x} = \lambda\boldsymbol{x}$
  - if we multiply both sides by $\bar{\boldsymbol{x}}^T$

$$\bar{\boldsymbol{x}}^T A\boldsymbol{x} = \bar{\boldsymbol{x}}^T(\lambda\boldsymbol{x}) = \lambda(\bar{\boldsymbol{x}}^T\boldsymbol{x}) = \lambda(\boldsymbol{x}\cdot\boldsymbol{x}) = \lambda\|\boldsymbol{x}\|^2$$

$$\lambda = \frac{\bar{\boldsymbol{x}}^T A\boldsymbol{x}}{\|\boldsymbol{x}\|^2}$$

  - Since the denominator is real, we prove that λ is real by showing

$$\overline{\bar{\boldsymbol{x}}^T A\boldsymbol{x}} = \bar{\boldsymbol{x}}^T A\boldsymbol{x}$$

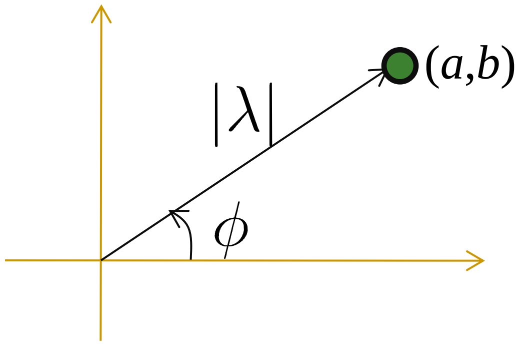
$$\overline{\bar{\boldsymbol{x}}^T A\boldsymbol{x}} = \bar{\bar{\boldsymbol{x}}}^T\overline{A\boldsymbol{x}} = \boldsymbol{x}^T\overline{A\boldsymbol{x}} = (\overline{A\boldsymbol{x}})^T\boldsymbol{x} = (\bar{A}\bar{\boldsymbol{x}})^T\boldsymbol{x} = (A\bar{\boldsymbol{x}})^T\boldsymbol{x} = \bar{\boldsymbol{x}}^T A^T\boldsymbol{x} = \bar{\boldsymbol{x}}^T A\boldsymbol{x}$$

Theorem 5.3.7

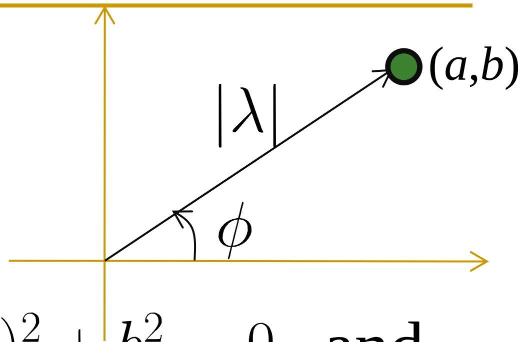
- Theorem 5.3.7: The eigenvalues of the real matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ are $\lambda = a \pm b_i$. If *a* and *b* are not both zero, then this matrix can be factored as

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

where ϕ is the angle from the positive *x*-axis to the ray that joins the origin to the point (*a*,*b*)



Proof of Theorem 5.3.7



- The characteristic equation of C is $(\lambda - a)^2 + b^2 = 0$, and the eigenvalues of C are $\lambda = a \pm b_i$
- Assuming that *a* and *b* are not both zero, let ϕ be the angle from the positive *x*-axis to the ray that joins the origin to the point (*a*,*b*). The angle ϕ is an argument of the eigenvalue $\lambda = a + b_i$, so we have

$$a = |\lambda| \cos\phi \qquad b = |\lambda| \sin\phi$$

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix} \begin{bmatrix} \frac{a}{|\lambda|} & -\frac{b}{|\lambda|} \\ \frac{b}{|\lambda|} & \frac{a}{|\lambda|} \end{bmatrix} = \begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

Theorem 5.3.8

- Let A be a real 2×2 matrix with complex eigenvalues $\lambda = a \pm b_i$ where $b \neq 0$. If **x** is an eigenvector of A corresponding to $\lambda = a - b_i$, then the matrix $P = [\mathrm{Re}(\boldsymbol{x}) \quad \mathrm{Im}(\boldsymbol{x})]$ is invertible and

$$A = P \begin{bmatrix} a & -b \\ b & a \end{bmatrix} P^{-1}$$

Example

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

- Factor the matrix using the eigenvalue $\lambda = -i$ and the corresponding eigenvector
- Solution: Let us denote the eigenvector that corresponds to $\lambda = -i$ by **x**.

$$a = 0 \quad b = 1 \quad \mathrm{Re}(\boldsymbol{x}) = \begin{bmatrix} -\frac{2}{5} \\ 1 \end{bmatrix} \quad \mathrm{Im}(\boldsymbol{x}) = \begin{bmatrix} -\frac{1}{5} \\ 0 \end{bmatrix}$$

- Thus,

$$P = [\mathrm{Re}(\boldsymbol{x}) \quad \mathrm{Im}(\boldsymbol{x})] = \begin{bmatrix} -\frac{2}{5} & -\frac{1}{5} \\ 1 & 0 \end{bmatrix}$$

- *A* can be factored as

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

Geometric Interpretation of Theorem 5.3.8

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix} \begin{bmatrix} \frac{a}{|\lambda|} & -\frac{b}{|\lambda|} \\ \frac{b}{|\lambda|} & \frac{a}{|\lambda|} \end{bmatrix} = \begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

- Let us denote the matrices on the right by S and R_ϕ
- Rewrite Theorem 5.3.8

$$A = PSR_\phi P^{-1} = P \begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} P^{-1}$$

- If we now view *P* as the transition matrix from the basis *B*=[Re(**x**) Im(**x**)] to the standard basis, then this equation tells us that computing a product $A\mathbf{x}_0$ can be broken into a three-step process

Geometric Interpretation of Theorem 5.3.8

- Step 1. Map $\mathbf{x}_0$ from standard coordinates into *B*-coordinates by forming the product $P^{-1}\mathbf{x}_0$.
- Step 2. Rotate and scale the vector $P^{-1}\mathbf{x}_0$ by forming the product $SR_\phi P^{-1}\boldsymbol{x}_0$
- Step 3. Map the rotated and scaled vector back to standard coordinates to obtain $A\boldsymbol{x}_0 = PSR_\phi P^{-1}\boldsymbol{x}_0$

Power Sequences

- If A is the standard matrix for an operator on R^n and $\mathbf{x}_0$ is some fixed factor in R^n, then one might be interested in the behavior of the power sequence $\mathbf{x}_0, A\mathbf{x}_0, A^2\mathbf{x}_0, \ldots, A^k\mathbf{x}_0, \ldots$
- With the help of MATLAB one can show that if the first 100 terms are plotted as ordered pairs (x,y), then the points move along the elliptical path show in Figure 5.3.4a
- To understand why, we need to examine the eigenvalues and eigenvectors of A.

Power Sequences

$$A = \begin{bmatrix} \frac{1}{2} & \frac{3}{4} \\ -\frac{3}{5} & \frac{11}{10} \end{bmatrix} \qquad \begin{aligned} \lambda_1 &= \tfrac{4}{5} - \tfrac{3}{5}i \qquad & \boldsymbol{v}_1 &= (\tfrac{1}{2} + i, 1) \\ \lambda_2 &= \tfrac{4}{5} + \tfrac{3}{5}i \qquad & \boldsymbol{v}_2 &= (\tfrac{1}{2} - i, 1) \end{aligned}$$

- We obtain the factorization

$$\begin{bmatrix} \frac{1}{2} & \frac{3}{4} \\ -\frac{3}{5} & \frac{11}{10} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}$$

$$A \quad = \quad P \qquad R_\phi \qquad P^{-1}$$

- R_ϕ is a rotation about the origin through the angle ϕ whose tangent is

$$\tan\phi = \frac{\sin\phi}{\cos\phi} = \frac{3/5}{4/5} = \frac{3}{4} \qquad \phi = \tan^{-1}\tfrac{3}{4} \approx 36.9^\circ$$

Power Sequences

- The matrix P is the transition matrix from the basis
$$B = \{\mathrm{Re}(\boldsymbol{x}), \mathrm{Im}(\boldsymbol{x})\} = \{(\tfrac{1}{2}, 1), (1, 0)\}$$
to the standard basis, and P^{-1} is the transition matrix from the standard basis to the basis B.
- If n is a positive integer,
$$A^n \boldsymbol{x}_0 = (P R_\phi P^{-1})^n \boldsymbol{x}_0 = P R_\phi^n P^{-1} \boldsymbol{x}_0$$
so the product $A^n\mathbf{x}_0$ can be compute by first mapping $\mathbf{x}_0$ into the point $P^{-1}\mathbf{x}_0$ in B-coordinates, then multiplying by R_ϕ^n to rotate this point about the origin through the angle $n\phi$, and then multiplying $R_\phi^n P^{-1} \boldsymbol{x}_0$ by P to map the resulting point back to the standard coordinates.

Power Sequences

- In B-coordinates each successive multiplication by A causes the point $P^{-1}\mathbf{x}_0$ to advance through an angle ϕ , thereby tracing a circular orbit about the origin.
- However, the basis is skewed (not orthogonal), so when the points on the circular orbit are transformed back to standard coordinates, the effect is to distort the circular orbit into the elliptical orbit traced by $A^n\mathbf{x}_0$.

Power Sequences

$$\begin{bmatrix} \frac{1}{2} & \frac{3}{4} \\ -\frac{3}{5} & \frac{11}{10} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

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

[$\mathbf{x}_0$ is mapped to *B*-coordinates]

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

[The point (1, ½) is rotated through the angle ϕ]

$$= \begin{bmatrix} \frac{5}{4} \\ \frac{1}{2} \end{bmatrix}$$

[The point (1, ½) is mapped to standard coordinates]