Eigenvalue Problems

Rafikul Alam
Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati - 781039, INDIA

Outline

- Eigenvalues and eigenvectors
- Multiplicity and diagonalizability
- Schur Theorem
- Eigenvector of triangular matrix
- Block Schur form
- Trace and eigenvalues

Eigenvalue problem

Let A be an $n \times n$ matrix. Then the eigenvalue problem for A is to solve

$$Av = \lambda v$$

for a nonzero vector v and a scalar λ. The scalar λ is called an eigenvalue of A and the vector v is called an eigenvector of A corresponding to λ.

Eigenvalue problem

Let A be an $n \times n$ matrix. Then the eigenvalue problem for A is to solve

$$Av = \lambda v$$

for a nonzero vector v and a scalar λ. The scalar λ is called an eigenvalue of A and the vector v is called an eigenvector of A corresponding to λ.

Example: Consider the matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$. Then 4 is an eigenvalue of A with $[1, 1]^\top$ is a corresponding eigenvector.

Eigenvalue problem

Let A be an $n \times n$ matrix. Then the eigenvalue problem for A is to solve

$$Av = \lambda v$$

for a nonzero vector v and a scalar λ. The scalar λ is called an eigenvalue of A and the vector v is called an eigenvector of A corresponding to λ.

Example: Consider the matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$. Then 4 is an eigenvalue of A with $[1, 1]^\top$ is a corresponding eigenvector.

Next, consider the rotation $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Then $Av \neq \lambda v$ for any $\lambda \in \mathbb{R}$ and $0 \neq v \in \mathbb{R}^2$. Hence we must consider eigenvalue problems over the field $\mathbb{C}$.

Eigenvalue problem

Let A be an $n \times n$ matrix. Then the eigenvalue problem for A is to solve

$$Av = \lambda v$$

for a nonzero vector v and a scalar λ. The scalar λ is called an eigenvalue of A and the vector v is called an eigenvector of A corresponding to λ.

Example: Consider the matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$. Then 4 is an eigenvalue of A with $[1, 1]^\top$ is a corresponding eigenvector.

Next, consider the rotation $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Then $Av \neq \lambda v$ for any $\lambda \in \mathbb{R}$ and $0 \neq v \in \mathbb{R}^2$. Hence we must consider eigenvalue problems over the field $\mathbb{C}$.

Eigenvalue problems arise in many applications. For instance, Google search engine solves an eigenvalue problem for ranking web pages. Computation of eigenvalues and eigenvectors of a matrix is a major task.

Eigenspace

Definition: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n\times n}$.

Eigenspace

Definition: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n \times n}$. Then $N(A - \lambda I) = \{\mathbf{x} : A\mathbf{x} = \lambda \mathbf{x}\}$ is called the eigenspace of A corresponding to λ. Note that the eigenspace of A corresponding to λ consists of all eigenvectors of A corresponding to λ together with the zero vector.

Eigenspace

Definition: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n\times n}$. Then $N(A - \lambda I) = \{\mathbf{x} : A\mathbf{x} = \lambda \mathbf{x}\}$ is called the eigenspace of A corresponding to λ. Note that the eigenspace of A corresponding to λ consists of all eigenvectors of A corresponding to λ together with the zero vector.

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then the following conditions are equivalent.

- $\lambda \in \mathbb{C}$ is an eigenvalue of A.

Eigenspace

Definition: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n\times n}$. Then $N(A - \lambda I) = \{\mathbf{x} : A\mathbf{x} = \lambda\mathbf{x}\}$ is called the eigenspace of A corresponding to λ. Note that the eigenspace of A corresponding to λ consists of all eigenvectors of A corresponding to λ together with the zero vector.

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then the following conditions are equivalent.

- $\lambda \in \mathbb{C}$ is an eigenvalue of A.
- $N(A - \lambda I)$ is a nontrivial subspace of $\mathbb{C}^n$.

Eigenspace

Definition: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n\times n}$. Then $N(A-\lambda I) = \{\mathbf{x} : A\mathbf{x} = \lambda\mathbf{x}\}$ is called the eigenspace of A corresponding to λ. Note that the eigenspace of A corresponding to λ consists of all eigenvectors of A corresponding to λ together with the zero vector.

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then the following conditions are equivalent.

- $\lambda \in \mathbb{C}$ is an eigenvalue of A.
- $N(A-\lambda I)$ is a nontrivial subspace of $\mathbb{C}^n$.
- $\mathrm{rank}(A-\lambda I) < n$.

Eigenspace

Definition: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n \times n}$. Then $N(A - \lambda I) = \{\mathbf{x} : A\mathbf{x} = \lambda \mathbf{x}\}$ is called the eigenspace of A corresponding to λ. Note that the eigenspace of A corresponding to λ consists of all eigenvectors of A corresponding to λ together with the zero vector.

Theorem: Let $A \in \mathbb{C}^{n \times n}$. Then the following conditions are equivalent.

- $\lambda \in \mathbb{C}$ is an eigenvalue of A.
- $N(A - \lambda I)$ is a nontrivial subspace of $\mathbb{C}^n$.
- $\mathrm{rank}(A - \lambda I) < n$.
- $\det(A - \lambda I) = 0$.

Eigenspace

Definition: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n\times n}$. Then $N(A - \lambda I) = \{\mathbf{x} : A\mathbf{x} = \lambda\mathbf{x}\}$ is called the eigenspace of A corresponding to λ. Note that the eigenspace of A corresponding to λ consists of all eigenvectors of A corresponding to λ together with the zero vector.

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then the following conditions are equivalent.

- $\lambda \in \mathbb{C}$ is an eigenvalue of A.
- $N(A - \lambda I)$ is a nontrivial subspace of $\mathbb{C}^n$.
- $\mathrm{rank}(A - \lambda I) < n$.
- $\det(A - \lambda I) = 0$.

Definition: Let A be an $n \times n$ matrix. Then

- $p(\lambda) := \det(\lambda I - A)$ is called the characteristic polynomial of A.

Eigenspace

Definition: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n\times n}$. Then $N(A - \lambda I) = \{\mathbf{x} : A\mathbf{x} = \lambda \mathbf{x}\}$ is called the eigenspace of A corresponding to λ. Note that the eigenspace of A corresponding to λ consists of all eigenvectors of A corresponding to λ together with the zero vector.

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then the following conditions are equivalent.

- $\lambda \in \mathbb{C}$ is an eigenvalue of A.
- $N(A - \lambda I)$ is a nontrivial subspace of $\mathbb{C}^n$.
- $\mathrm{rank}(A - \lambda I) < n$.
- $\det(A - \lambda I) = 0$.

Definition: Let A be an $n \times n$ matrix. Then

- $p(\lambda) := \det(\lambda I - A)$ is called the characteristic polynomial of A.
- $\det(\lambda I - A) = 0$ is called the characteristic equation of A.

Eigenspace

Definition: Let λ be an eigenvalue of a matrix $A \in \mathbb{C}^{n\times n}$. Then $N(A - \lambda I) = \{\mathbf{x} : A\mathbf{x} = \lambda\mathbf{x}\}$ is called the eigenspace of A corresponding to λ. Note that the eigenspace of A corresponding to λ consists of all eigenvectors of A corresponding to λ together with the zero vector.

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then the following conditions are equivalent.

- $\lambda \in \mathbb{C}$ is an eigenvalue of A.
- $N(A - \lambda I)$ is a nontrivial subspace of $\mathbb{C}^n$.
- $\mathrm{rank}(A - \lambda I) < n$.
- $\det(A - \lambda I) = 0$.

Definition: Let A be an $n \times n$ matrix. Then

- $p(\lambda) := \det(\lambda I - A)$ is called the characteristic polynomial of A.
- $\det(\lambda I - A) = 0$ is called the characteristic equation of A.

Definition: The set $\Lambda(A) := \{\lambda \in \mathbb{C} : \mathrm{rank}(A - \lambda I) < n\}$ is called the spectrum of A.

Eigenvalues

- The eigenvalues of A are the roots of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$. Hence A has at most n eigenvalues.

Eigenvalues

- The eigenvalues of A are the roots of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$. Hence A has at most n eigenvalues.
- The eigenvalues of a triangular matrix are its diagonal entries.

Eigenvalues

- The eigenvalues of A are the roots of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$. Hence A has at most n eigenvalues.
- The eigenvalues of a triangular matrix are its diagonal entries.
- A is invertible iff 0 is not an eigenvalue of A.

Eigenvalues

- The eigenvalues of A are the roots of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$. Hence A has at most n eigenvalues.
- The eigenvalues of a triangular matrix are its diagonal entries.
- A is invertible iff 0 is not an eigenvalue of A.

Example: Let $A := \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$. Then we have

$$\det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{bmatrix} = 0 \Longrightarrow \quad \lambda = 4, -2.$$

Eigenvalues

- The eigenvalues of A are the roots of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$. Hence A has at most n eigenvalues.
- The eigenvalues of a triangular matrix are its diagonal entries.
- A is invertible iff 0 is not an eigenvalue of A.

Example: Let $A := \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$. Then we have

$$\det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{bmatrix} = 0 \Longrightarrow \quad \lambda = 4, -2.$$

$$N(A - 4I) = \left\{ \mathbf{x} : \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix} \mathbf{x} = \mathbf{0} \right\} = \text{span}([1, 1]^\top),$$

Eigenvalues

- The eigenvalues of A are the roots of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$. Hence A has at most n eigenvalues.
- The eigenvalues of a triangular matrix are its diagonal entries.
- A is invertible iff 0 is not an eigenvalue of A. => det(A - 0*I) = 0 => det(A) = 0. So, A is singular, hence not invertible.

Example: Let $A := \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$. Then we have

$$\det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{bmatrix} = 0 \implies \quad \lambda = 4, -2.$$

$$N(A - 4I) = \left\{ \mathbf{x} : \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix} \mathbf{x} = \mathbf{0} \right\} = \operatorname{span}([1, 1]^\top),$$

$$N(A + 2I) = \left\{ \mathbf{x} : \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \mathbf{x} = \mathbf{0} \right\} = \operatorname{span}([1, -1]^\top).$$

Algebraic and geometric multiplicities

Definition: Let λ be an eigenvalue of $A \in \mathbb{C}^{n\times n}$.

Algebraic and geometric multiplicities

Definition: Let λ be an eigenvalue of $A \in \mathbb{C}^{n\times n}$.

- The algebraic multiplicity of λ is the multiplicity of λ as a root of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$.

Algebraic and geometric multiplicities

Definition: Let λ be an eigenvalue of $A \in \mathbb{C}^{n\times n}$.

- The algebraic multiplicity of λ is the multiplicity of λ as a root of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$.
- The geometric multiplicity of λ is the dimension of $N(A - \lambda I)$.

Algebraic and geometric multiplicities

Definition: Let λ be an eigenvalue of $A \in \mathbb{C}^{n\times n}$.

- The algebraic multiplicity of λ is the multiplicity of λ as a root of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$.
- The geometric multiplicity of λ is the dimension of $N(A - \lambda I)$.

Theorem: Let λ be an eigenvalue of A. Then

$$\text{geometric multiplicity of } \lambda \leq \text{ algebraic multiplicity of } \lambda.$$

Algebraic and geometric multiplicities

Definition: Let λ be an eigenvalue of $A \in \mathbb{C}^{n\times n}$.

- The algebraic multiplicity of λ is the multiplicity of λ as a root of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$.
- The geometric multiplicity of λ is the dimension of $N(A - \lambda I)$.

Theorem: Let λ be an eigenvalue of A. Then

$$\text{geometric multiplicity of } \lambda \leq \text{ algebraic multiplicity of } \lambda.$$

Example: Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Then 1 is the eigenvalue of A of algebraic multiplicity 3 and geometric multiplicity 1.

Algebraic and geometric multiplicities

Definition: Let λ be an eigenvalue of $A \in \mathbb{C}^{n\times n}$.

- The algebraic multiplicity of λ is the multiplicity of λ as a root of the characteristic polynomial $p(\lambda) := \det(\lambda I - A)$.
- The geometric multiplicity of λ is the dimension of $N(A - \lambda I)$.

Theorem: Let λ be an eigenvalue of A. Then

$$\text{geometric multiplicity of } \lambda \leq \text{ algebraic multiplicity of } \lambda.$$

Example: Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Then 1 is the eigenvalue of A of algebraic multiplicity 3 and geometric multiplicity 1. Indeed, we have

$$N(A - I) = N\left(\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) \Longrightarrow \dim(N(A - I)) = 1.$$

Eigenvalues via characteristic polynomial

Let $A \in \mathbb{C}^{n\times n}$ and $p(x) := \det(xI - A)$ be the characteristic polynomial of A. Then $\lambda \in \mathbb{C}$ is an eigenvalue of $A \iff p(\lambda) = 0$. Thus eigenvalues of A can be computed by finding the roots of $p(x)$. To see the efficacy of this method, consider $A := \mathrm{diag}(1, 2, \ldots, 22)$.

Eigenvalues via characteristic polynomial

Let $A \in \mathbb{C}^{n\times n}$ and $p(x) := \det(xI - A)$ be the characteristic polynomial of A. Then $\lambda \in \mathbb{C}$ is an eigenvalue of $A \iff p(\lambda) = 0$. Thus eigenvalues of A can be computed by finding the roots of $p(x)$. To see the efficacy of this method, consider $A := \mathrm{diag}(1, 2, \ldots, 22)$.

```
>> A = diag(1:22) ; % eigenvalues 1,2,...,22
>> p = charpoly(A); % coefficents of the poly p(x)
>> rt = roots(p); % roots of p(x)
>> rt(6:9) % displays four roots
```

Eigenvalues via characteristic polynomial

Let $A \in \mathbb{C}^{n\times n}$ and $p(x) := \det(xI - A)$ be the characteristic polynomial of A. Then $\lambda \in \mathbb{C}$ is an eigenvalue of $A \iff p(\lambda) = 0$. Thus eigenvalues of A can be computed by finding the roots of $p(x)$. To see the efficacy of this method, consider $A := \mathrm{diag}(1, 2, \ldots, 22)$.

```
>> A = diag(1:22) ; % eigenvalues 1,2,...,22
>> p = charpoly(A); % coeffcents of the poly p(x)
>> rt = roots(p); % roots of p(x)
>> rt(6:9) % displays four roots
ans =
17.564435730130054 + 0.661474607910510i
17.564435730130054 - 0.661474607910510i
15.388471193084563 + 0.581348923952655i
15.388471193084563 - 0.581348923952655i
```

These roots do not have even a single correct digit! Never compute roots of $\det(\lambda I - A)$ for computing eigenvalues of A.

Similar matrices

The diagonal and triangular matrices are nice in the sense that their eigenvalues are displayed on the diagonals.

Similar matrices

The diagonal and triangular matrices are nice in the sense that their eigenvalues are displayed on the diagonals.

It would be desirable to transform a square matrix to a diagonal or triangular matrix in such a way that they have exactly the same eigenvalues.

Similar matrices

The diagonal and triangular matrices are nice in the sense that their eigenvalues are displayed on the diagonals.

It would be desirable to transform a square matrix to a diagonal or triangular matrix in such a way that they have exactly the same eigenvalues.

Gaussian elimination reduces a matrix to a triangular matrix. Unfortunately, this process does not preserve the eigenvalues.

Similar matrices

The diagonal and triangular matrices are nice in the sense that their eigenvalues are displayed on the diagonals.

It would be desirable to transform a square matrix to a diagonal or triangular matrix in such a way that they have exactly the same eigenvalues.

Gaussian elimination reduces a matrix to a triangular matrix. Unfortunately, this process does not preserve the eigenvalues.

Definition: Let $A, B \in \mathbb{C}^{n\times n}$. Then A is said to be similar to B if there is an invertible matrix $P \in \mathbb{C}^{n\times n}$ such that $P^{-1}AP = B$.

Similar matrices

The diagonal and triangular matrices are nice in the sense that their eigenvalues are displayed on the diagonals.

It would be desirable to transform a square matrix to a diagonal or triangular matrix in such a way that they have exactly the same eigenvalues.

Gaussian elimination reduces a matrix to a triangular matrix. Unfortunately, this process does not preserve the eigenvalues.

Definition: Let $A, B \in \mathbb{C}^{n \times n}$. Then A is said to be similar to B if there is an invertible matrix $P \in \mathbb{C}^{n \times n}$ such that $P^{-1}AP = B$.

The map $A \longmapsto P^{-1}AP$ is called a similarity transformation of A.

Similar matrices

The diagonal and triangular matrices are nice in the sense that their eigenvalues are displayed on the diagonals.

It would be desirable to transform a square matrix to a diagonal or triangular matrix in such a way that they have exactly the same eigenvalues.

Gaussian elimination reduces a matrix to a triangular matrix. Unfortunately, this process does not preserve the eigenvalues.

Definition: Let $A, B \in \mathbb{C}^{n \times n}$. Then A is said to be similar to B if there is an invertible matrix $P \in \mathbb{C}^{n \times n}$ such that $P^{-1}AP = B$.

The map $A \longmapsto P^{-1}AP$ is called a similarity transformation of A.

Note that $\det(P^{-1}AP - \lambda I) = \det(A - \lambda I)$. Consequently, similar matrices have the same characteristic polynomials and hence have the same eigenvalues.

Proof: https://math.stackexchange.com/questions/8339/similar-matrices-have-the-same-eigenvalues-with-the-same-geometric-multiplicity

Diagonalization

Note that A and $B := P^{-1}AP$ have the same eigenvalues. Indeed, we have $\det(\lambda I - A) = \det(\lambda I - B)$. Moreover, $B\mathbf{v} = \lambda\mathbf{v} \iff A\mathbf{u} = \lambda\mathbf{u}$, where $\mathbf{u} := P\mathbf{v}$.

Diagonalization

Note that A and $B := P^{-1}AP$ have the same eigenvalues. Indeed, we have $\det(\lambda I - A) = \det(\lambda I - B)$. Moreover, $B\mathbf{v} = \lambda\mathbf{v} \iff A\mathbf{u} = \lambda\mathbf{u}$, where $\mathbf{u} := P\mathbf{v}$.

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$. Then $AP = PB$, where $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. ■

Diagonalization

Note that A and $B := P^{-1}AP$ have the same eigenvalues. Indeed, we have $\det(\lambda I - A) = \det(\lambda I - B)$. Moreover, $B\mathbf{v} = \lambda\mathbf{v} \iff A\mathbf{u} = \lambda\mathbf{u}$, where $\mathbf{u} := P\mathbf{v}$.

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$. Then $AP = PB$, where $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. ■

Definition: An matrix $A \in \mathbb{C}^{n\times n}$ is said to be diagonalizable if there is an invertible matrix $P \in \mathbb{C}^{n\times n}$ such that $D := P^{-1}AP$ is a diagonal matrix.

Diagonalization

Note that A and $B := P^{-1}AP$ have the same eigenvalues. Indeed, we have $\det(\lambda I - A) = \det(\lambda I - B)$. Moreover, $B\mathbf{v} = \lambda\mathbf{v} \iff A\mathbf{u} = \lambda\mathbf{u}$, where $\mathbf{u} := P\mathbf{v}$.

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$. Then $AP = PB$, where $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. $\blacksquare$

Definition: An matrix $A \in \mathbb{C}^{n\times n}$ is said to be diagonalizable if there is an invertible matrix $P \in \mathbb{C}^{n\times n}$ such that $D := P^{-1}AP$ is a diagonal matrix.

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then A is diagonalizable $\iff$ A has n linearly independent eigenvectors.

Diagonalization

Note that A and $B := P^{-1}AP$ have the same eigenvalues. Indeed, we have $\det(\lambda I - A) = \det(\lambda I - B)$. Moreover, $B\mathbf{v} = \lambda\mathbf{v} \iff A\mathbf{u} = \lambda\mathbf{u}$, where $\mathbf{u} := P\mathbf{v}$.

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$. Then $AP = PB$, where $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. $\blacksquare$

Definition: An matrix $A \in \mathbb{C}^{n\times n}$ is said to be diagonalizable if there is an invertible matrix $P \in \mathbb{C}^{n\times n}$ such that $D := P^{-1}AP$ is a diagonal matrix.

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then A is diagonalizable $\iff$ A has n linearly independent eigenvectors. In particular, if A has n distinct eigenvalues then A is diagonalizable.

Proof: If $P^{-1}AP = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ then $A\mathbf{v}_j = \lambda_j\mathbf{v}_j$, where $\mathbf{v}_j := Pe_j$.

Diagonalization

Note that A and $B := P^{-1}AP$ have the same eigenvalues. Indeed, we have $\det(\lambda I - A) = \det(\lambda I - B)$. Moreover, $B\mathbf{v} = \lambda \mathbf{v} \iff A\mathbf{u} = \lambda \mathbf{u}$, where $\mathbf{u} := P\mathbf{v}$.

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$. Then $AP = PB$, where $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. ■

Definition: An matrix $A \in \mathbb{C}^{n\times n}$ is said to be diagonalizable if there is an invertible matrix $P \in \mathbb{C}^{n\times n}$ such that $D := P^{-1}AP$ is a diagonal matrix.

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then A is diagonalizable $\iff$ A has n linearly independent eigenvectors. In particular, if A has n distinct eigenvalues then A is diagonalizable.

Proof: If $P^{-1}AP = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ then $A\mathbf{v}_j = \lambda_j \mathbf{v}_j$, where $\underline{\mathbf{v}_j := Pe_j}$. Conversely, if $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for $i = 1 : n$ and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent then As uj = ej

$$A[\mathbf{v}_1, \ldots, \mathbf{v}_n] = [\mathbf{v}_1, \ldots, \mathbf{v}_n]\mathrm{diag}(\lambda_1, \ldots, \lambda_n). \blacksquare$$

Schur triangularization theorem

Schur Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then there is a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that U^*AU is upper triangular.

Schur triangularization theorem

Schur Theorem: Let $A \in \mathbb{C}^{n \times n}$. Then there is a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that $U^* AU$ is upper triangular.

Proof: Use induction on n. The result holds for $n = 1$. Assume that the result holds for $n - 1$.

Schur triangularization theorem

Schur Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then there is a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that U^*AU is upper triangular.

Proof: Use induction on n. The result holds for $n = 1$. Assume that the result holds for $n-1$.

Let $A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$ be such that $\|\mathbf{u}_1\|_2 = 1$. Choose $\mathbf{u}_2, \ldots, \mathbf{u}_n$ such that $U_1 := \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n\end{bmatrix}$ is unitary.

Schur triangularization theorem

Schur Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then there is a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that U^*AU is upper triangular.

Proof: Use induction on n. The result holds for $n = 1$. Assume that the result holds for $n-1$.

Let $A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$ be such that $\|\mathbf{u}_1\|_2 = 1$. Choose $\mathbf{u}_2, \ldots, \mathbf{u}_n$ such that $U_1 := \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n\end{bmatrix}$ is unitary. Then $AU_1 = \begin{bmatrix}\lambda_1\mathbf{u}_1 & A\mathbf{u}_2 & \cdots & A\mathbf{u}_n\end{bmatrix}$ and hence

$$U_1^*AU_1 = \begin{bmatrix} \lambda_1 & \mathbf{h}^\top \\ 0 & H \end{bmatrix} \text{ for some } \mathbf{h} \in \mathbb{C}^{n-1} \text{ and } H \in \mathbb{C}^{(n-1)\times(n-1)}.$$

Schur triangularization theorem

Schur Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then there is a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that U^*AU is upper triangular.

Proof: Use induction on n. The result holds for $n = 1$. Assume that the result holds for $n-1$.

Let $A\mathbf{u}_1 = \lambda_1 \mathbf{u}_1$ be such that $\|\mathbf{u}_1\|_2 = 1$. Choose $\mathbf{u}_2, \ldots, \mathbf{u}_n$ such that $U_1 := \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix}$ is unitary. Then $AU_1 = \begin{bmatrix} \lambda_1 \mathbf{u}_1 & A\mathbf{u}_2 & \cdots & A\mathbf{u}_n \end{bmatrix}$ and hence

$$U_1^* A U_1 = \begin{bmatrix} \lambda_1 & \mathbf{h}^\top \\ 0 & H \end{bmatrix} \text{ for some } \mathbf{h} \in \mathbb{C}^{n-1} \text{ and } H \in \mathbb{C}^{(n-1)\times(n-1)}.$$

By induction hypothesis, there exists a unitary V such that $V^*HV = T$ is upper triangular.

Schur triangularization theorem

Schur Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then there is a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that U^*AU is upper triangular.

Proof: Use induction on n. The result holds for $n = 1$. Assume that the result holds for $n-1$.

Let $A\mathbf{u}_1 = \lambda_1 \mathbf{u}_1$ be such that $\|\mathbf{u}_1\|_2 = 1$. Choose $\mathbf{u}_2, \ldots, \mathbf{u}_n$ such that
$U_1 := \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix}$ is unitary. Then $AU_1 = \begin{bmatrix} \lambda_1 \mathbf{u}_1 & A\mathbf{u}_2 & \cdots & A\mathbf{u}_n \end{bmatrix}$ and hence

$$U_1^* A U_1 = \begin{bmatrix} \lambda_1 & \mathbf{h}^\top \\ 0 & H \end{bmatrix} \text{ for some } \mathbf{h} \in \mathbb{C}^{n-1} \text{ and } H \in \mathbb{C}^{(n-1)\times(n-1)}.$$

By induction hypothesis, there exists a unitary V such that $V^*HV = T$ is upper triangular.
Set $U := U_1 \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix}$.

Schur triangularization theorem

Schur Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then there is a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that U^*AU is upper triangular.

Proof: Use induction on n. The result holds for $n = 1$. Assume that the result holds for $n-1$.

Let $A\mathbf{u}_1 = \lambda_1 \mathbf{u}_1$ be such that $\|\mathbf{u}_1\|_2 = 1$. Choose $\mathbf{u}_2, \ldots, \mathbf{u}_n$ such that $U_1 := \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix}$ is unitary. Then $AU_1 = \begin{bmatrix} \lambda_1\mathbf{u}_1 & A\mathbf{u}_2 & \cdots & A\mathbf{u}_n \end{bmatrix}$ and hence

$$U_1^* A U_1 = \begin{bmatrix} \lambda_1 & \mathbf{h}^\top \\ 0 & H \end{bmatrix} \text{ for some } \mathbf{h} \in \mathbb{C}^{n-1} \text{ and } H \in \mathbb{C}^{(n-1)\times(n-1)}.$$

By induction hypothesis, there exists a unitary V such that $V^*HV = T$ is upper triangular. Set $U := U_1 \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix}$. Then

$$U^*AU = \begin{bmatrix} \lambda_1 & \mathbf{h}^\top V \\ 0 & T \end{bmatrix} = \text{ upper triangular. } \blacksquare$$

Spectral theorem for Hermitian matrices

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then A is Hermitian $\iff$ there exist a unitary matrix U and a real diagonal matrix D such that
$$U^*AU = D.$$

Spectral theorem for Hermitian matrices

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then A is Hermitian $\iff$ there exist a unitary matrix U and a real diagonal matrix D such that

$$U^*AU = D.$$

Thus a Hermitian matrix is unitarily diagonalizable and has n orthonormal eigenvectors.

Spectral theorem for Hermitian matrices

Theorem: Let $A \in \mathbb{C}^{n \times n}$. Then A is Hermitian $\iff$ there exist a unitary matrix U and a real diagonal matrix D such that

$$U^* A U = D.$$

Thus a Hermitian matrix is unitarily diagonalizable and has n orthonormal eigenvectors.

Proof: By Schur theorem, there is a unitary matrix U and an upper triangular matrix D such that $U^* A U = D$.

Spectral theorem for Hermitian matrices

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then A is Hermitian $\iff$ there exist a unitary matrix U and a real diagonal matrix D such that

$$U^*AU = D.$$

Thus a Hermitian matrix is unitarily diagonalizable and has n orthonormal eigenvectors.

Proof: By Schur theorem, there is a unitary matrix U and an upper triangular matrix D such that $U^*AU = D$.

Now $A^* = A \iff D^* = D \iff D$ is a diagonal matrix with real diagonals.

Spectral theorem for Hermitian matrices

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then A is Hermitian $\iff$ there exist a unitary matrix U and a real diagonal matrix D such that

$$U^*AU = D.$$

Thus a Hermitian matrix is unitarily diagonalizable and has n orthonormal eigenvectors.

Proof: By Schur theorem, there is a unitary matrix U and an upper triangular matrix D such that $U^*AU = D$.

Now $A^* = A \iff D^* = D \iff D$ is a diagonal matrix with real diagonals. If $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $U = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ then

Spectral theorem for Hermitian matrices

Theorem: Let $A \in \mathbb{C}^{n\times n}$. Then A is Hermitian $\iff$ there exist a unitary matrix U and a real diagonal matrix D such that

$$U^*AU = D.$$

Thus a Hermitian matrix is unitarily diagonalizable and has n orthonormal eigenvectors.

Proof: By Schur theorem, there is a unitary matrix U and an upper triangular matrix D such that $U^*AU = D$.

Now $A^* = A \iff D^* = D \iff D$ is a diagonal matrix with real diagonals. If $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $U = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ then

$$A\mathbf{u}_j = \lambda_j \mathbf{u}_j \text{ for } j = 1 : n$$

Spectral theorem for Hermitian matrices

Theorem: Let $A \in \mathbb{C}^{n \times n}$. Then A is Hermitian $\iff$ there exist a unitary matrix U and a real diagonal matrix D such that

$$U^* A U = D.$$

$\boxed{\text{Thus a Hermitian matrix is unitarily diagonalizable and has } n \text{ orthonormal eigenvectors.}}$

Proof: By Schur theorem, there is a unitary matrix U and an upper triangular matrix D such that $U^* A U = D$.

Now $A^* = A \iff D^* = D \iff D$ is a diagonal matrix with real diagonals. If $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $U = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ then

$$A\mathbf{u}_j = \lambda_j \mathbf{u}_j \text{ for } j = 1 : n$$

and $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are orthonormal. ■

Eigenvectors of triangular matrix

By Schur's theorem, an $n \times n$ matrix can be transformed into triangular (Schur) form by unitary similarity and eigenvalues of a triangular matrix are the diagonal entries.

Eigenvectors of triangular matrix

By Schur's theorem, an $n \times n$ matrix can be transformed into triangular (Schur) form by unitary similarity and eigenvalues of a triangular matrix are the diagonal entries.

Eigenvectors of a triangular matrix are less obvious, but still easy to compute. Suppose that A is upper triangular and λ is a simple eigenvalue of A.

Eigenvectors of triangular matrix

By Schur's theorem, an $n \times n$ matrix can be transformed into triangular (Schur) form by unitary similarity and eigenvalues of a triangular matrix are the diagonal entries.

Eigenvectors of a triangular matrix are less obvious, but still easy to compute. Suppose that A is upper triangular and λ is a simple eigenvalue of A. Then $A - \lambda I$ will be of the form

$$A - \lambda I = \begin{bmatrix} A_{11} & \mathbf{b} & A_{13} \\ & 0 & \mathbf{h}^\top \\ & & A_{33} \end{bmatrix}.$$

Eigenvectors of triangular matrix

By Schur's theorem, an $n \times n$ matrix can be transformed into triangular (Schur) form by unitary similarity and eigenvalues of a triangular matrix are the diagonal entries.

Eigenvectors of a triangular matrix are less obvious, but still easy to compute. Suppose that A is upper triangular and λ is a simple eigenvalue of A. Then $A - \lambda I$ will be of the form

$$A - \lambda I = \begin{bmatrix} A_{11} & \mathbf{b} & A_{13} \\ & 0 & \mathbf{h}^\top \\ & & A_{33} \end{bmatrix}.$$

Then solving the upper triangular system $A_{11}\mathbf{x} = \mathbf{b}$ and defining

$$\mathbf{v} := \begin{bmatrix} \mathbf{x} \\ -1 \\ \mathbf{0} \end{bmatrix}$$

it follows that $\mathbf{v}$ is an eigenvector of A corresspónding to λ, that is, $A\mathbf{v} = \lambda \mathbf{v}$.

Block upper triangular form

Real Schur Theorem: Let $A \in \mathbb{R}^{n\times n}$. Then there is an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ such that $U^* AU$ is block upper triangular given by

Block upper triangular form

Real Schur Theorem: Let $A \in \mathbb{R}^{n\times n}$. Then there is an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ such that U^*AU is block upper triangular given by

$$U^*AU = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1m} \\ & T_{22} & \cdots & T_{2m} \\ & & \ddots & \vdots \\ & & & T_{mm} \end{bmatrix},$$

where $T_{jj} = \lambda_j \in \mathbb{R}$ or $T_{jj} = \begin{bmatrix} \alpha_j & -\beta_j \\ \beta_j & \alpha_j \end{bmatrix}$ with $\alpha_j, \beta_j \in \mathbb{R}$ for $j = 1 : m$.

Block upper triangular form

Real Schur Theorem: Let $A \in \mathbb{R}^{n\times n}$. Then there is an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ such that U^*AU is block upper triangular given by

$$U^*AU = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1m} \\ & T_{22} & \cdots & T_{2m} \\ & & \ddots & \vdots \\ & & & T_{mm} \end{bmatrix},$$

where $T_{jj} = \lambda_j \in \mathbb{R}$ or $T_{jj} = \begin{bmatrix} \alpha_j & -\beta_j \\ \beta_j & \alpha_j \end{bmatrix}$ with $\alpha_j, \beta_j \in \mathbb{R}$ for $j = 1 : m$.

Further, each 2×2 block $T_{jj} = \begin{bmatrix} \alpha_j & -\beta_j \\ \beta_j & \alpha_j \end{bmatrix}$ corresponds to a pair of complex conjugate eigenvalues $\alpha_j \pm i\beta_j$ of A.

Trace and eigenvalues

Theorem: Let $A \in \mathbb{C}^{n\times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ (counted according to their algebraic multiplicities).

Trace and eigenvalues

Theorem: Let $A \in \mathbb{C}^{n\times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ (counted according to their algebraic multiplicities). Then

$$\mathrm{Trace}(A) = \lambda_1 + \cdots + \lambda_n \text{ and } \det(A) = \prod_{j=1}^{n} \lambda_j.$$

Proof: By Schur theorem, there is a unitary matrix U such that $T := U^*AU$ is upper triangular.

Trace and eigenvalues

Theorem: Let $A \in \mathbb{C}^{n\times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ (counted according to their algebraic multiplicities). Then

$$\mathrm{Trace}(A) = \lambda_1 + \cdots + \lambda_n \text{ and } \det(A) = \prod_{j=1}^{n} \lambda_j.$$

Proof: By Schur theorem, there is a unitary matrix U such that $T := U^* A U$ is upper triangular. Hence

$$\det(A) = \det(T) = \prod_{j=1}^{n} \lambda_j.$$

Trace and eigenvalues

Theorem: Let $A \in \mathbb{C}^{n\times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ (counted according to their algebraic multiplicities). Then

$$\mathrm{Trace}(A) = \lambda_1 + \cdots + \lambda_n \text{ and } \det(A) = \prod_{j=1}^{n} \lambda_j.$$

Proof: By Schur theorem, there is a unitary matrix U such that $T := U^*AU$ is upper triangular. Hence

$$\det(A) = \det(T) = \prod_{j=1}^{n} \lambda_j.$$

Now $\mathrm{Trace}(A) = \mathrm{Trace}(UTU^*) = \mathrm{Trace}(T) = \lambda_1 + \cdots + \lambda_n$. ■

Exercise: Let $A, B \in \mathbb{C}^{n\times n}$. Show that $\mathrm{Trace}(AB) = \mathrm{Trace}(BA)$.