ERG 2012B
Advanced Engineering Mathematics II

Part II: Linear Algebra

Lecture #15
Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

If $\mathbf{A}=[a_{jk}]$ is a given n x n matrix, consider the vector equation:

$$\mathbf{Ax} = \lambda\mathbf{x}$$

where λ is a number. It is clear that the zero vector $\mathbf{x}=\mathbf{0}$ is a solution to this for any value of λ.

- A value of λ for which there is a solution $\mathbf{x}\neq\mathbf{0}$ is called an **eigenvalue** or **characteristic value** (or *latent root*)
- The corresponding solutions $\mathbf{x}\neq\mathbf{0}$ themselves are called **eigenvectors** or **characteristic vectors** of **A**
- The set of eigenvectors is called the **spectrum of A**.
- The largest of the absolute values of the eigenvalues is called the **spectral radius of A**
- The set of all eigenvectors corresponding to an eigenvalue of **A,** together with **0,** form a vector space - the **eigenspace.**

Example 1

Determination of eigenvalues and eigenvectors

Find the eigenvalues and eigenvectors of the matrix

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

Solution: eigenvalues. These **must** be determined **first**.

$$\mathbf{Ax} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} x_a \\ x_b \end{bmatrix} = \lambda \begin{bmatrix} x_a \\ x_b \end{bmatrix}$$

or

$$\begin{aligned} -5x_a + 2x_b &= \lambda x_a \\ 2x_a - 2x_b &= \lambda x_b \end{aligned}$$

so that:

$$\begin{aligned} (-5-\lambda)x_a + \quad 2x_b &= 0 \\ 2x_a \quad + (-2-\lambda)x_b &= 0 \end{aligned}$$

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

which is homogeneous. By Cramer's rule it has solution $\mathbf{x} \neq \mathbf{0}$ iff its coefficient determinant is zero.

Example 1

Determination of eigenvalues and eigenvectors

Find the eigenvalues and eigenvectors of the matrix

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

Solution: eigenvalues.

$$D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix} = (-5-\lambda)(-2-\lambda)-4 = \lambda^2+7\lambda+6 = 0$$

We call $D(\lambda)$ the **characteristic determinant** or, if expanded, the **characteristic polynomial**, and $D(\lambda)=0$ the **characteristic equation** of **A**.

The solutions of this quadratic equation are $\lambda_1=-1$ and $\lambda_2=-6$. These are the eigenvalues of **A**

Example 1

Determination of eigenvalues and eigenvectors

Find the eigenvalues and eigenvectors of the matrix

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

Solution: eigenvector of A corresponding to λ_1.

This vector is found by setting $\lambda=\lambda_1=-1$ in the original equations

$$\begin{aligned} (-5-\lambda)x_a + \quad 2x_b \quad &= 0 \qquad \text{or} \qquad -4x_a + 2x_b = 0 \\ 2x_a \quad + (-2-\lambda)x_b &= 0 \qquad\qquad\quad\; 2x_a - \; x_b = 0 \end{aligned}$$

i.e. $x_b=2x_a$ where x_a is chosen arbitrarily. Let $x_a=1$ then $x_b=2$ and an eigenvector corresponding to $\lambda_1=-1$ is

$$\mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{Check:} \quad \mathbf{A}\mathbf{x}_1 = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \lambda\begin{bmatrix} -1 \\ -2 \end{bmatrix} = -1\mathbf{x_1} = \lambda_1\mathbf{x_1}$$

Example 1

Determination of eigenvalues and eigenvectors

Find the eigenvalues and eigenvectors of the matrix

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

Solution: eigenvector of A corresponding to λ_2.

Set $\lambda=\lambda_2=-6$ in the original equations

$$\begin{aligned} (-5-\lambda)x_a + \quad 2x_b \quad &= 0 \qquad \text{or} \qquad x_a + 2x_b = 0 \\ 2x_a \quad + (-2-\lambda)x_b &= 0 \qquad\qquad 2x_a + 4x_b = 0 \end{aligned}$$

i.e. $x_b=-x_a/2$ where x_a is chosen arbitrarily. Let $x_a=2$ then $x_b=-1$ and an eigenvector corresponding to $\lambda_2=-6$ is

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

Eigenvalues

Theorem 1 The eigenvalues of a square matrix **A** are the roots of the corresponding characteristic equation.

Hence an n x n matrix has at least one eigenvalue and at most n numerically different eigenvalues

The eigen*values* **must** be determined first. Once these are known, corresponding eigen*vectors* are obtained, for instance using Gauss elimination.

Theorem 2 If $\mathbf{x}$ is an eigenvector of a matrix $\mathbf{A}$ corresponding to an eigenvalue λ, so is $k\mathbf{x}$ with any $k \neq 0$

Proof: $\mathbf{Ax} = \lambda\mathbf{x}$ implies $k(\mathbf{Ax}) = \mathbf{A}(k\mathbf{x}) = \lambda(k\mathbf{x})$

Example 2

Multiple Eigenvalues

Find the eigenvalues and eigenvectors of the matrix

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

Solution The characteristic determinant gives the characteristic equation: $D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = -\lambda^3 - \lambda^2 + 21\lambda + 45 = 0$

$$\begin{vmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{vmatrix} = (-2-\lambda)[(1-\lambda)(-\lambda)-12] - 2(-2\lambda-6) - 3(-4+1-\lambda)$$

$$= (-2-\lambda)(-\lambda+\lambda^2-12) + (4\lambda+12) + (9+3\lambda\lambda$$

$$= 2\lambda - 2\lambda^2 + 24 + \lambda^2 - \lambda^3 + 12\lambda + 7\lambda + 21$$

$$= -\lambda^3 - \lambda^2 + 21\lambda + 45$$

Example 2

Multiple Eigenvalues

Find the eigenvalues and eigenvectors of the matrix

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

Solution The characteristic determinant gives the characteristic equation: $-\lambda^3 - \lambda^2 + 21\lambda + 45 = 0$

The roots (eigenvalues) are $\lambda_1=5$, $\lambda_2= \lambda_3=-3$

To find the eigenvectors use Gauss elimination on $(\mathbf{A}- \lambda\mathbf{I})\mathbf{x}=\mathbf{0}$ letting $\lambda=\mathbf{5}$ first:

$$\begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix} \Rightarrow \begin{bmatrix} -7 & 2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$x_a=(2x_b-3x_c)/7$

$x_b=-2x_c$

$$\Rightarrow \mathbf{x_1} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Example 2

Multiple Eigenvalues

Find the eigenvalues and eigenvectors of the matrix

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

Solution The characteristic determinant gives the characteristic equation: $-\lambda^3 - \lambda^2 + 21\lambda + 45 = 0$

The roots (eigenvalues) are $\lambda_1=5$, $\lambda_2= \lambda_3=-3$

To find the eigenvectors use Gauss elimination on $(\mathbf{A}- \lambda\mathbf{I})\mathbf{x}=\mathbf{0}$ letting $\lambda = \mathbf{-3}$:

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

$x_a=-(2x_b-3x_c)$

let $x_b=1$; $x_c=0$

then let $x_b=0$; $x_c=1$

$$\Rightarrow \mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$$

Multiplicity of Eigenvalues

Multiple Eigenvalues

If an eigenvalue λ of a matrix **A** is a root of order M_λ of the characteristic polynomial, then M_λ is called the **algebraic multiplicity** of λ

The **geometric multiplicity** m_λ of λ, is defined to be the number of linearly independent eigenvectors corresponding to λ, thus, the dimension of the corresponding eigenspace.

Since the characteristic polynomial has degree n, the sum of all algebraic multiplicities equals n.

In example 2 for λ=-3 we have $m_\lambda = M_\lambda = 2$. In general $m_\lambda \leq M_\lambda$ which we can demonstrate as follows:

Example 3

The characteristic equation of the matrix:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{is} \quad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2 = 0$$

Hence $\lambda = 0$ is an eigenvalue of algebraic multiplicity 2

But its geometric multiplicity is only 1

Because eigenvectors result from:

$$-0x_a + x_b = 0$$

hence $x_b = 0$ in the form $\begin{bmatrix} x_a \\ 0 \end{bmatrix}$ i.e. only one independent vector

Example 4

Real matrices with complex eigenvalues and eigenvectors

Since real polynomials may have complex roots, a real matrix may have complex eigenvalues and eigenvectors.

E.g the characteristic equation of the skew-symmetric matrix:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{is} \quad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0$$

gives the eigenvalues $\lambda_1 = i$, $\lambda_2 = -i$.

Eigenvectors are obtained from $-ix_a + x_b = 0$ and $ix_a + x_b = 0$ respectively, and choosing $x_a = 1$ we get:

$$\begin{bmatrix} 1 \\ i \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

Symmetric, Skew-Symmetric and Orthogonal Matrices

A **symmetric** matrix is a square matrix such that

$$\mathbf{A}^{\mathrm{T}} = \mathbf{A} \qquad \text{thus } a_{kj} = a_{jk}$$

A **skew-symmetric** matrix is a square matrix such that

$$\mathbf{A}^{\mathrm{T}} = -\mathbf{A} \qquad \text{thus } a_{kj} = -a_{jk}$$

An **orthogonal** matrix is a square matrix such that

$$\mathbf{A}^{\mathrm{T}} = \mathbf{A}^{-1}$$

Example 1: the matrices

$$\underbrace{\begin{bmatrix} -3 & 1 & 5 \\ 1 & 0 & -2 \\ 5 & -2 & 4 \end{bmatrix}}_{\text{symmetric}}, \quad \underbrace{\begin{bmatrix} 0 & 9 & -12 \\ -9 & 0 & 20 \\ 12 & -20 & 0 \end{bmatrix}}_{\text{skew-symmetric}}, \quad \underbrace{\begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}}_{\text{orthogonal}}$$

Every skew-symmetric matrix has all main diagonal entries zero

Properties

Any real square matrix **A** may be written as the sum of a symmetric matrix **R** and a skew-symmetric matrix **S**, where

$$\mathbf{R} = \tfrac{1}{2}(\mathbf{A} + \mathbf{A}^{\mathbf{T}}) \quad \text{and} \quad \mathbf{S} = \tfrac{1}{2}(\mathbf{A} - \mathbf{A}^{\mathbf{T}})$$

$$\mathbf{R}^{\mathbf{T}} = \tfrac{1}{2}(\mathbf{A}^{\mathbf{T}} + \mathbf{A}) = \mathbf{R}$$

and $$\mathbf{S}^{\mathbf{T}} = \tfrac{1}{2}(\mathbf{A}^{\mathbf{T}} - \mathbf{A}) = -\mathbf{S}$$

Example 2:

$$\mathbf{A} = \begin{bmatrix} 3 & -4 & -1 \\ 6 & 0 & -1 \\ -3 & 13 & -4 \end{bmatrix} = \mathbf{R} + \mathbf{S} = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 6 \\ -2 & 6 & -4 \end{bmatrix} + \begin{bmatrix} 0 & -5 & 1 \\ 5 & 0 & -7 \\ -1 & 7 & 0 \end{bmatrix}$$

Theorem 1

Theorem 1:

Eigenvalues of symmetric and skew-symmetric matrices

(a) The eigenvalues of a symmetric matrix are real

(b) The eigenvalues of a skew-symmetric matrix are pure imaginary or zero.

Example 3:

$$\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \quad \text{so that} \quad \begin{vmatrix} 5-\lambda & 3 \\ 3 & 5-\lambda \end{vmatrix} = (5-\lambda)^2 - 9 = 0$$

$$\Rightarrow \lambda_1 = 8, \qquad \lambda_2 = 2 \quad \text{both are real}$$

Example 4:

$$\mathbf{A} = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix} \quad \text{so that} \quad \begin{vmatrix} -\lambda & 3 \\ -3 & -\lambda \end{vmatrix} = \lambda^2 + 9 = 0$$

$$\Rightarrow \lambda_1 = 3i, \qquad \lambda_2 = -3i \quad \text{both pure imaginary}$$

Orthogonal Transformations

These are transformations

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad \text{with } \mathbf{A} \text{ an orthogonal matrix}$$

With each vector $\mathbf{x}$ in R^n such a transformation assigns a vector $\mathbf{y}$ also in R^n. For instance, the plane rotation through θ

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is an orthogonal transformation.

It can be shown that any orthogonal transformation in the plane or in 3D space is a rotation (possibly combined with a reflection in a straight line or plane, respectively)

Invariance of Inner Product

Theorem 2: An orthogonal transformation preserves the value of the inner product of vectors

$$\mathbf{a} \bullet \mathbf{b} = \mathbf{a}^{\mathrm{T}}\mathbf{b}$$

Hence also the **length** or **norm** of a vector in R^n given by

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \bullet \mathbf{a}} = \sqrt{\mathbf{a}^{\mathrm{T}}\mathbf{a}}$$

Proof: Let $\mathbf{u} = \mathbf{Aa}$ and $\mathbf{v} = \mathbf{Ab}$ where $\mathbf{A}$ is orthogonal. We must show that $\mathbf{u}^{\mathrm{T}}\mathbf{v} = \mathbf{a}^{\mathrm{T}}\mathbf{b}$

We know that $\mathbf{u}^{\mathrm{T}} = (\mathbf{Aa})^{\mathrm{T}} = \mathbf{a}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$

Also $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ as $\mathbf{A}$ is orthogonal

Therefore: $\mathbf{u}^{\mathrm{T}}\mathbf{v} = (\mathbf{Aa})^{\mathrm{T}}\mathbf{Ab} = \mathbf{a}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{Ab} = \mathbf{a}^{\mathrm{T}}\mathbf{Ib} = \mathbf{a}^{\mathrm{T}}\mathbf{b}$

Orthonormality

Theorem 3: Orthonormality of column and row vectors

A real square matrix is orthogonal iff its column vectors (and also its row vectors) form an **orthogonal system**, i.e.

$$\mathbf{a}_j \bullet \mathbf{a}_k = \mathbf{a}_j^{\mathrm{T}} \mathbf{a}_k = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases} \qquad (*)$$

Proof: (a) Let **A** be orthogonal. Then $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}$, in terms of column vectors

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}} \\ \mathbf{a}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{a}_n^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}}\mathbf{a}_1 & \mathbf{a}_1^{\mathrm{T}}\mathbf{a}_2 & \cdots & \mathbf{a}_1^{\mathrm{T}}\mathbf{a}_n \\ \mathbf{a}_2^{\mathrm{T}}\mathbf{a}_1 & \mathbf{a}_2^{\mathrm{T}}\mathbf{a}_2 & \cdots & \mathbf{a}_2^{\mathrm{T}}\mathbf{a}_n \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_n^{\mathrm{T}}\mathbf{a}_1 & \mathbf{a}_n^{\mathrm{T}}\mathbf{a}_2 & \cdots & \mathbf{a}_n^{\mathrm{T}}\mathbf{a}_n \end{bmatrix} = \mathbf{I}$$

by the definition of **I** this implies that (*) is correct

Orthonormality

Theorem 3: Orthonormality of column and row vectors

A real square matrix is orthogonal iff its column vectors (and also its row vectors) form an **orthogonal system**, i.e.

$$\mathbf{a}_j \bullet \mathbf{a}_k = \mathbf{a}_j^{\mathrm{T}} \mathbf{a}_k = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases} \qquad (*)$$

Proof: (b) Conversely, if the column vectors of **A** satisfy (*) the off-diagonal entries are 0 and the diagonals are 1.

Hence $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}$ and $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{I}$. This implies $\mathbf{A}^{\mathrm{T}} = \mathbf{A}^{-1}$ and so **A** is orthogonal.

Determinant of Orthogonal Matrix

Theorem 4: The determinant of an orthogonal matrix has the value +1 or -1

Proof: This follows from

$$\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B} \quad \text{and}$$
$$\det \mathbf{A}^{\mathbf{T}} = \det \mathbf{A}$$

If $\mathbf{A}$ is orthogonal then:

$$1 = \det \mathbf{I} = \det (\mathbf{AA}^{-1}) = \det (\mathbf{AA}^{\mathbf{T}}) = \det \mathbf{A} \det \mathbf{A}^{\mathbf{T}} = (\det \mathbf{A})^2$$

Theorem 5: Eigenvalues of an orthogonal matrix. The eigenvalues of an orthogonal matrix are real or complex conjugates in pairs and have absolute value 1.

Proof: The 1st part is true for any real matrix as its characteristic polynomial has real coefficients. The 2nd part will be proved later.

Example 5

The orthogonal matrix

$$\begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$$

Has the characteristic equation:

$$-\lambda^3 + 2\lambda^2/3 + 2\lambda/3 - 1 = 0$$

One of the eigenvalues must be real (why?), hence +1 or –1. Trying –1 shows that it satisfies the equation.

Dividing by $(\lambda+1)$ gives $\lambda^2 - 5\lambda/3 + 1 = 0$ and the other two eigenvalues are:

$$(5+i\sqrt{11})/6 \text{ and } (5-i\sqrt{11})/6$$

Hermitian, Skew-Hermitian & Unitary Matrices

We introduce three classes of *complex* square matrices that generalize the three classes of *real* matrices just considered.

They have important applications, e.g. in quantum mechanics and systems theory.

We use the standard notation

$$\overline{\mathbf{A}} = \left[\overline{a}_{jk}\right]$$

the matrix replacing each entry of **A** by its complex conjugate and

$$\overline{\mathbf{A}}^{\mathrm{T}} = \left[\overline{a}_{kj}\right]$$

for the conjugate transpose. For example:

$$\mathbf{A} = \begin{bmatrix} 3+4i & -5i \\ -7 & 6-2i \end{bmatrix}, \quad \text{then} \quad \overline{\mathbf{A}}^{\mathrm{T}} = \begin{bmatrix} 3-4i & -7 \\ 5i & 6+2i \end{bmatrix}$$

Hermitian, Skew-Hermitian & Unitary Matrices

Definition: A square matrix $\mathbf{A} = [a_{jk}]$ is called

Hermitian if $\overline{\mathbf{A}}^{\mathrm{T}} = \mathbf{A}$ that is, $\overline{a}_{kj} = a_{jk}$

skew-Hermitian if $\overline{\mathbf{A}}^{\mathrm{T}} = -\mathbf{A}$ that is, $\overline{a}_{kj} = -a_{jk}$

Unitary if $\overline{\mathbf{A}}^{\mathrm{T}} = \mathbf{A}^{-1}$

From these definitions we see the following:

- If **A** is Hermitian, the entries on the main diagonal $\overline{a}_{jj} = a_{jj}$
- If **A** is skew-Hermitian, then $\overline{a}_{jj} = -a_{jj}$
  that is, if we set a_{jj} = x+iy this means x-iy=-(x+iy) i.e. x=0 and a_{jj} is pure imaginary or 0

Example 1

Hermitian, skew-Hermitian and Unitary matrices

$$\mathbf{A} = \begin{bmatrix} 4 & 1-3i \\ 1+3i & 7 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 3i & 2+i \\ -2+i & -i \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} \frac{1}{2}i & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2}i \end{bmatrix}$$

Hermitian — skew-Hermitian — Unitary

If a Hermitian matrix is real then $\overline{\mathbf{A}}^{\mathbf{T}} = \mathbf{A}^{\mathbf{T}} = \mathbf{A}$ Hence a real Hermitian matrix is symmetric.

Similarly, if a skew-Hermitian matrix is real then $\overline{\mathbf{A}}^{\mathbf{T}} = \mathbf{A}^{\mathbf{T}} = -\mathbf{A}$ Hence it is a skew-symmetric matrix.

Finally, if a unitary matrix is real then $\overline{\mathbf{A}}^{\mathbf{T}} = \mathbf{A}^{\mathbf{T}} = \mathbf{A}^{-1}$ Hence it is orthogonal.

This shows that these are just generalizations of symmetric, skew-symmetric and orthogonal matrices.

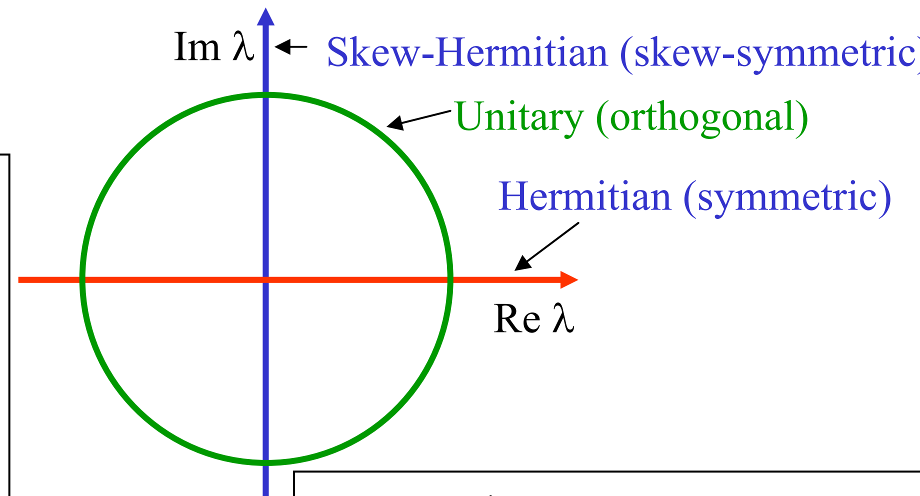
Theorem 1

Eigenvalues

(a) The eigenvalues of a Hermitian matrix (and hence a symmetric matrix) are real

(b) The eigenvalues of a skew-Hermitian matrix (and hence a skew-symmetric matrix) are pure imaginary or zero

(c) The eigenvalues of a unitary matrix (and hence an orthogonal matrix) have absolute value 1



Matrix	characteristic equation	Eigenvalues
A	$\lambda^2-11\lambda+18=0$	9,2
B	$\lambda^2-2i\lambda+8=0$	4i, -2i
C	$\lambda^2-i\lambda-1=0$	$\pm(\sqrt{3}+i)/2$

and $|\pm(\sqrt{3}+i)/2|^2=(3+1)/4=1$

$$\mathbf{A}=\begin{bmatrix} 4 & 1-3i \\ 1+3i & 7 \end{bmatrix},\quad \mathbf{B}=\begin{bmatrix} 3i & 2+i \\ -2+i & -i \end{bmatrix},\quad \mathbf{C}=\begin{bmatrix} \frac{1}{2}i & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2}i \end{bmatrix}$$

Proof - Theorem 1

Let λ be an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ a corresponding eigenvector

Then: $$\mathbf{Ax} = \lambda\mathbf{x}$$

(a) The eigenvalues of a Hermitian matrix are real

Let $\mathbf{A}$ be Hermitian. Now

$$\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ax} = \overline{\mathbf{x}}^{\mathbf{T}}\lambda\mathbf{x} = \lambda\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{x}$$

Now $\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{x} = \overline{x}_1 x_1 + \ldots. + \overline{x}_n x_n = |x_1|^2 + \ldots. + |x_n|^2$ is real, and is not 0 since $\mathbf{x}\neq\mathbf{0}$. Hence we may divide to get

$$\lambda = \frac{\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ax}}{\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{x}}$$

λ is real if the numerator is real. This is true if it is equal to its conjugate. As the numerator is a number – not a vector or a matrix – transposition does not effect it, so, using Hermitian properties

$$\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ax} = (\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ax})^{\mathbf{T}} = \mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\overline{\mathbf{x}} = \mathbf{x}^{\mathbf{T}}\overline{\mathbf{A}}\overline{\mathbf{x}} = \overline{(\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ax})}$$

Proof - Theorem 1

Let λ be an eigenvalue of **A** and **x** a corresponding eigenvector

Then: $\mathbf{Ax} = \lambda\mathbf{x}$

(b) The eigenvalues of a skew-Hermitian matrix are pure imaginary or zero. In this case the argument is the same but:

$$\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ax} = (\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ax})^{\mathbf{T}} = \mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\overline{\mathbf{x}} = -\mathbf{x}^{\mathbf{T}}\overline{\mathbf{A}}\overline{\mathbf{x}} = -\overline{(\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ax})}$$

$$\lambda = \frac{\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ax}}{\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{x}}$$

In other words λ is a complex number that equals minus its complex conjugate, a+*i*b=-(a-*i*b). Hence a=0 so that λ is pure imaginary or zero.

Proof - Theorem 1

Let λ be an eigenvalue of **A** and **x** a corresponding eigenvector

Then: $$\mathbf{Ax} = \lambda\mathbf{x}$$

(c) The eigenvalues of a unitary matrix have absolute value 1.

Now if **A** is unitary then using the conjugate transpose we have:

$$\mathbf{Ax} = \lambda\mathbf{x} \quad \text{and} \quad (\overline{\mathbf{A}}\overline{\mathbf{x}})^{\mathbf{T}} = (\overline{\lambda}\overline{\mathbf{x}})^{\mathbf{T}} = \overline{\lambda}\overline{\mathbf{x}}^{\mathbf{T}}$$

Multiplying the two left sides and the two right sides

But **A** is unitary so that:

$$(\overline{\mathbf{A}}\overline{\mathbf{x}})^{\mathbf{T}}\mathbf{Ax} = \overline{\mathbf{x}}^{\mathbf{T}}\overline{\mathbf{A}}^{\mathbf{T}}\mathbf{Ax} = \overline{\mathbf{x}}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{Ax} = \overline{\mathbf{x}}^{\mathbf{T}}\mathbf{Ix} = \overline{\mathbf{x}}^{\mathbf{T}}\mathbf{x}$$

combining the two equations $\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{x} = |\lambda|^2\,\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{x}$

Now divide by $\overline{\mathbf{x}}^{\mathbf{T}}\mathbf{x}\ (\neq 0)$ to get $|\lambda|^2 = 1$

Forms

The numerator used in the proof of a) and b) $\bar{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ is called a **form** in the components x_1, ...,x_n of **x** and **A** is called its *coefficient matrix*. When n=2 we get:

$$\bar{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \begin{bmatrix} \bar{x}_1 & \bar{x}_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \bar{x}_1 & \bar{x}_2 \end{bmatrix} \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

$$= a_{11}\bar{x}_1 x_1 + a_{12}\bar{x}_1 x_2 + a_{21}\bar{x}_2 x_1 + a_{22}\bar{x}_2 x_2$$

In general

$$\begin{aligned} \bar{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}\bar{x}_j x_k &= a_{11}\bar{x}_1 x_1 + \ldots + a_{1n}\bar{x}_1 x_n \\ &+ a_{21}\bar{x}_2 x_1 + \ldots + a_{2n}\bar{x}_2 x_n \\ &+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &+ a_{n1}\bar{x}_n x_1 + \ldots + a_{nn}\bar{x}_n x_n \end{aligned}$$

Forms

If **x** and **A** are real then this becomes:

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k = a_{11}x_1^2 + a_{12}x_1x_2 \ldots + a_{1n}x_1x_n$$

$$+ a_{21}x_2x_1 + a_{22}x_2^2 \ldots + a_{2n}x_2x_n$$

$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$+ a_{n1}x_nx_1 + a_{n2}x_nx_2 \ldots + a_{nn}x_n^2$$

and is called a **quadratic form**. Without restriction we may then assume that coefficient matrix to be symmetric, because we can take off diagonals together in pairs and then write the result as a sum of two equal terms – illustrated in the example

Example 3

Quadratic Form. Symmetric coefficient matrix C

Let

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1^2 + 4x_1x_2 + 6x_2x_1 + 2x_2^2$$

$$= 3x_1^2 + 10x_1x_2 + 2x_2^2$$

Here 4 + 6 = 10 and so does 5+5. We can make a corresponding symmetric matrix C=[c_{jk}], where c_{jk}=½(a_{jk}+a_{kj}), thus c_{11}=3, c_{12}=c_{21}=5, c_{22}=2

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1^2 + 5x_1x_2 + 5x_2x_1 + 2x_2^2$$

$$= 3x_1^2 + 10x_1x_2 + 2x_2^2$$

Hermitian Forms

If the matrix **A** is Hermitian or skew-Hermitian the form is called a **Hermitian form** or **skew-Hermitian form**, respectively. These forms have the following property, which makes them important in physics.

Theorem 1* For every choice of the vector **x** the value of a Hermitian form is real, and the value of a skew-Hermitian form is pure imaginary or zero.

Proof: The previous proof assumed **x** to be an eigenvector, but the proof remains valid for any vectors.

Example 4

Hermitian Form

If

$$\mathbf{A}=\begin{bmatrix}3 & 2-i\\ 2+i & 4\end{bmatrix} \text{ and } \quad \mathbf{x}=\begin{bmatrix}1+i\\ 2i\end{bmatrix}$$

then

$$\overline{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\mathbf{x}=\begin{bmatrix}1-i & -2i\end{bmatrix}\begin{bmatrix}3 & 2-i\\ 2+i & 4\end{bmatrix}\begin{bmatrix}1+i\\ 2i\end{bmatrix}$$

$$=\begin{bmatrix}1-i & -2i\end{bmatrix}\begin{bmatrix}3(1+i)+(2-i)2i\\ (2+i)(1+i)+4\cdot 2i\end{bmatrix}=34$$