CHAPTER 8

Linear Algebra: Matrix Eigenvalue Problems

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

Matrix eigenvalue problem:

Given the vector equation

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

with **A** is a given square matrix, determine the unknown scalar λ and unknown vector **x**.

- Trivial solution $\mathbf{x} = \mathbf{0}$ not interesting – consider solutions with $\mathbf{x} \neq \mathbf{0}$.
- There can be more than one solution.

Some Terminologies:

- λ is called an **eigenvalue** or *characteristic value,* a *latent root* of the matrix **A**. ("Eigen" is German and means "proper" or "characteristic.")
- The corresponding solutions $\mathbf{x} \neq \mathbf{0}$ of (1) are called the **eigenvectors** or *characteristic vectors* of **A** corresponding to that eigenvalue λ.
- The set of all eigenvalues of **A** is called the **spectrum** of **A**.
- The spectrum consists of at least one eigenvalue and at most of n numerically different eigenvalues (later).
- The largest of the absolute values of the eigenvalues of **A** is called the *spectral radius* of **A** (later).

Determination of Eigenvalues and Eigenvectors

EXAMPLE 1: To illustrate the steps in terms of the matrix

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

Solution:

(a) Finding Eigenvalues (which are to be determined **first**).

From (1) $\quad \mathbf{Ax} = \lambda\mathbf{x}$

(2) $\qquad \mathbf{Ax} - \lambda\mathbf{x} = \mathbf{0}$

(3) $\qquad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ $\qquad$ *homogeneous* equation!

- Not interested in situation with trivial solution $\mathbf{x}=\mathbf{0}$.
- From Section 7.6, the homogeneous equation (3) has nontrivial solution iff rank $(\mathbf{A}-\lambda\mathbf{I}) < 2$, i.e., det $(\mathbf{A}-\lambda\mathbf{I}) = 0$.

EXAMPLE 1 (continued)

- This yields

$$D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix}$$

$$(4^*) \qquad = (-5-\lambda)(-2-\lambda) - 4 = \lambda^2 + 7\lambda + 6 = 0.$$

- $D(\lambda)$ is the **characteristic determinant** or, if expanded, the **characteristic polynomial**.

- $D(\lambda) = 0$ is called the **characteristic equation** of **A**.

- The solutions of the **characteristic equation** of **A**, $D(\lambda) = 0$, are the eigenvalues of **A.**

- In this case, solution of (4*) are $\lambda_1 = -1$ and $\lambda_2 = -6$ (eigenvalues determined).

EXAMPLE 1 (continued)

(b) Finding Eigenvector

(i) Eigenvector of A corresponding to λ_1. Substituting $\lambda = \lambda_1 = -1$ in (3*):

$$(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{x} = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

- With Rank $(\mathbf{A} - \lambda_1 \mathbf{I}) = 1$, procedures of solving homogeneous equation in Section 7.6 → nontrivial solution with 1 free parameter.
- Choosing x_2 as free parameter and letting $x_2 = 2$, we have $x_1 = 0.5x_2 = 1$. Hence the eigenvector for λ_1:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \text{ Check: } \mathbf{A}\mathbf{x}_1 = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = (-1)\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \lambda_1 \mathbf{x}_1$$

- Eigenvector $\mathbf{x}_1$ is determined up to a scalar multiple.

EXAMPLE 1 (continued)

(ii) Eigenvector of A corresponding to λ_2. Substituting $\lambda = \lambda_2 = -6$ in (2*):

$$(\mathbf{A} - \lambda_2 \mathbf{I})\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

- With Rank $(\mathbf{A} - \lambda_2 \mathbf{I}) = 1$, procedures of solving homogeneous equation in Section 7.6 → nontrivial solution with 1 free parameter.
- Choosing x_2 as free parameter and letting $x_2 = -1$, we have $x_1 = -2x_2 = 2$. Hence the eigenvector for λ_2:

$$\mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \text{Check:} \quad \mathbf{A}\mathbf{x}_2 = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -12 \\ 6 \end{bmatrix} = (-6)\mathbf{x}_2 = \lambda_2 \mathbf{x}_2$$

- Eigenvector $\mathbf{x}_2$ is determined up to a scalar multiple.

EXAMPLE (Opening example in Sec. 8.1)

The matrix equation:

$$\mathbf{Ax} = \lambda \mathbf{x}, \mathbf{A} = \begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}$$

- The characteristic equation is
  $$\lambda^2 - 13\lambda + 30 = (\lambda - 10)(\lambda - 3) = 0.$$
- Eigenvalues are $\{10,\ 3\}$.
- Corresponding eigenvectors are $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$,
  respectively, determined up to a scalar multiple.

Previous examples concern matrix with distinct (different) eigenvalues. There are also matrix that produces multiple (or repeated) eigenvalues.

EXAMPLE 2 Matrix with Multiple Eigenvalues

Find the eigenvalues and eigenvectors of the *n* x *n* matrix

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

Solution.

- Conduct characteristic determinant to yield characteristic equation det($\mathbf{A} - \lambda\,\mathbf{I}$)=0

$$\Longrightarrow \quad -\lambda^3 - \lambda^2 + 21\lambda + 45 = 0.$$

- The roots of the characteristic equation yield eigenvalues of $\mathbf{A}$: $\lambda_1 = 5,\ \underbrace{\lambda_2 = \lambda_3 = -3}_{\text{Eigenvalue repeated 2 times, or eigenvalue with multiplicity 2}}.$

EXAMPLE 2 (continued)

- Eigenvector of **A** corresponding to $\lambda_1 = 5$. Substituting $\lambda = \lambda_1 = 5$ in (2*):

$$(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{x} = \begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

  – Apply elementary row operations (as in Section 7.6):

$$(\mathbf{A} - \lambda_1 \mathbf{I}) = \begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix} \longrightarrow \begin{bmatrix} -7 & 2 & -3 \\ 0 & -24/7 & -48/7 \\ 0 & 0 & 0 \end{bmatrix}.$$

  – Hence, rank $(\mathbf{A} - \lambda_1 \mathbf{I})=2$ → nontrivial solution with $(n\text{-}2)=1$ free parameter.

  – Choosing x_3 as free parameter and letting $x_3 = 1$, we have the eigenvector for λ_1 as:

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

Determined up to a scalar *c*

EXAMPLE 2 (continued)

- Eigenvector of **A** corresponding to $\lambda_2 = \lambda_3 = 5$. Substituting $\lambda_{2,3} = -3$ in (2*):

$$(\mathbf{A} - \lambda_{2,3}\mathbf{I})\mathbf{x} = (\mathbf{A} + 3\mathbf{I}) = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

  - Apply elementary row operations (as in Section 7.6):

$$(\mathbf{A} - \lambda_{2,3}\mathbf{I}) = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

  - Hence, rank $(\mathbf{A} - \lambda_{2,3}\mathbf{I})=1$ → homogeneous equations above has nontrivial solution with $(n\text{-}1)=2$ free parameter.

EXAMPLE 2 (continued)

- Picking x_2 and x_3 as free parameters, we have

$$x_1 + 2x_2 - 3x_3 = 0 \rightarrow x_1 = -2x_2 + 3x_3.$$

- Hence, eigenvectors for λ_2 and λ_3 are generally given by:

$$\mathbf{x}_{2,3} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$$

- Choosing $x_2 = 1$, $x_3 = 0$ and $x_2 = 0$, $x_3 = 1$, we obtain two linearly independent eigenvectors of $\mathbf{A}$ corresponding to $\lambda = -3$, which we can assign to λ_2 and λ_3:

$$\mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_3 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

SUMMARY: Determining Eigenvalues/Eigenvectors

- Generally, eigenvalues is a set of n linear equations in n unknowns:

(1) $$\mathbf{Ax} = \lambda \mathbf{x} \quad \longleftrightarrow \quad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = \lambda x_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = \lambda x_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = \lambda x_n. \end{cases}$$

- Transferring the terms on the right side to the left side, we have

(2) $$(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0} \quad \longleftrightarrow \quad \begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn} - \lambda)x_n = 0. \end{cases}$$

- Homogeneous linear system (2) has a nonzero solution iff corresponding determinant of the coefficient matrix is zero:

$$(4) \qquad D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

(**Characteristic polynomial** of **A** – nth degree polynomial of λ)

- Eigenvalues of **A**, $\lambda_1, \lambda_2, \ldots, \lambda_n$ obtained as roots of the Characteristic polynomial.
- Corresponding eigenvector $\mathbf{x}_i$ for λ_i obtained by substituting λ_i into (2) and solving for $\mathbf{x}_i$.

- Specifically, for distinct eigenvalue λ_i , i.e., λ_i appearing only once among the eigenvalues of **A**, substituting λ_i yields rank $(\mathbf{A} - \lambda_i \mathbf{I})=(n\text{-}1)$

  → homogeneous equations $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x}=\mathbf{0}$ has nontrivial solution with $(n\text{-}(n\text{-}1))=1$ free parameter

- On the other hand, for eigenvalue λ_i with multiplicity p, i.e., λ_i appearing p times among the eigenvalues of **A**, substituting λ_i yields rank $(\mathbf{A} - \lambda_i \mathbf{I})=(n\text{-}p)$

  → homogeneous equations $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x}=\mathbf{0}$ has nontrivial solution with $(n\text{-}(n\text{-}p))=p$ free parameters.

  → by choosing the free parameters properly, we produce p linearly independent eigenvectors assigned to the p eigenvalues of λ_i .

Note:

- Determine eigenvalues first and then eigenvectors.
- If **x** is eigenvector of λ, then c**x** is also an eigenvector of λ, where c is any scalar. An eigenvector **x** is determined only up to a constant factor.
- We can also choose to **normalize x**, that is, multiply it by a scalar to get a unit vector.

Theorem 1

Eigenvalues

- The eigenvalues of a square matrix **A** are the roots of the characteristic equation (4) of **A**.
- An $n \times n$ matrix has at least one eigenvalue (of multiplicity n) and at most n numerically distinct eigenvalues.

Theorem 2

Eigenvectors and Eigenspace

- If **w** and **x** are eigenvectors of a matrix **A** corresponding to ***the same*** eigenvalue λ, so are $\mathbf{w} + \mathbf{x}$ (provided $\mathbf{x} \neq -\mathbf{w}$) and $k\mathbf{x}$ for any $k \neq 0$.
- The eigenvectors corresponding to one and the same eigenvalue λ *of* **A**, together with **0**, hence form a vector space, called the **eigenspace** of **A** corresponding to that λ.

EXAMPLE 2 revisited

The n x n matrix

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

has eigenvalues $\lambda_1 = 5$, $\lambda_2 = \lambda_3 = -3$.

- For eigenvalue $\lambda_1 = 5$, the span of eigenvector $\mathbf{x}_1$, together with the trivial solution **0**, form a vector space, which is the eigenspace of λ_1.

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

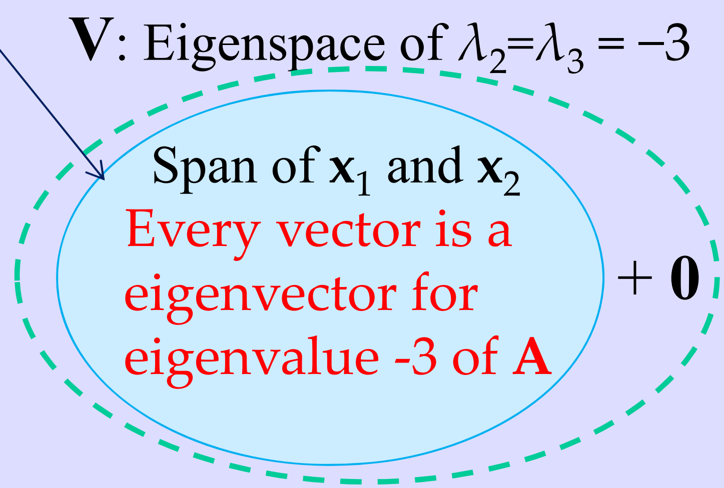
V: Eigenspace of λ_1

Span of $\mathbf{x}_1$:
Every vector is a eigenvector of λ_1

$+\ \mathbf{0}$

- For eigenvalue $\lambda_2 = \lambda_3 = -3$, the span of eigenvector $\mathbf{x}_2$ and $\mathbf{x}_3$, together with the trivial solution **0,** form a vector space, which is the eigenspace of the eigenvalue $= -3$ of **A.**

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



Theorem 3

Eigenvalues of the Transpose

The transpose $\mathbf{A}^T$ of a square matrix $\mathbf{A}$ has the same eigenvalues as $\mathbf{A}$.

- Why? Noting that det ($\mathbf{M}$) = det ($\mathbf{M}^T$), one has

$$\det(\mathbf{A}-\lambda\mathbf{I}) = \det[(\mathbf{A}-\lambda\mathbf{I})^T],$$

or

$$\det(\mathbf{A}-\lambda\mathbf{I}) = \det(\mathbf{A}^T-\lambda\mathbf{I}).$$

This means that if λ is an eigenvalue of $\mathbf{A}$, then λ is an eigenvalue of $\mathbf{A}^T$ as well.

- However, $\mathbf{A}^T$ and $\mathbf{A}$ generally have different eigenvectors!

EXAMPLE:

Matrix $\mathbf{A} = \begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}$ has

- Eigenvalues $\{10, \ 3\}$.
- Eigenvectors are $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, respectively.

(up to a scalar multiple)

On the other hand, matrix $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 6 & 4 \\ 3 & 7 \end{bmatrix}$ has

- Same eigenvalues of $\{10, \ 3\}$.
- But different eigenvectors of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ -3 \end{bmatrix}$, respectively.

(up to a scalar multiple)

8.2 Some Applications of Eigenvalue Problems

EXAMPLE 1 Stretching of an Elastic Membrane

- An elastic membrane in the x_1x_2-plane with boundary circle $x_1^2 + x_2^2 = 1$ is stretched so that point P: (x_1, x_2) becomes point Q: (y_1, y_2).
- The relation between them (Fig. 160) is:

$$(1) \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{A}\mathbf{x} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix};$$

- Find the **principal directions** – the directions of the position vector **x** for which the direction of the position vector **y** is the same or exactly opposite.
- What is the new shape of the boundary of the circle under the deformation?

EXAMPLE 1 (continued 1)
Stretching of an Elastic Membrane

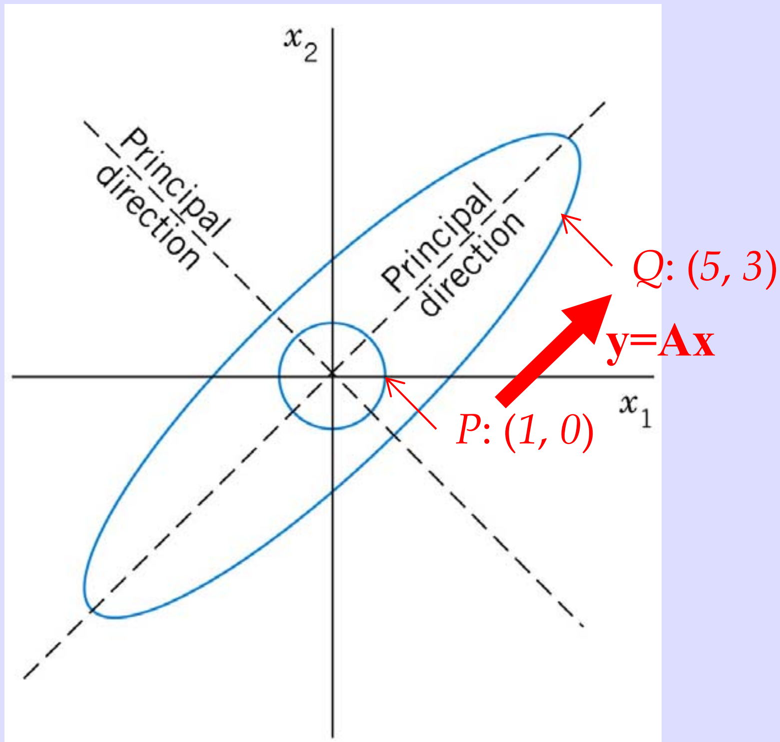


Fig. 160. Undeformed and deformed membrane in Example 1

EXAMPLE 1 (continued 2)
Stretching of an Elastic Membrane

Solution.

- We desire vectors $\mathbf{x}$ such that $\mathbf{y} = \mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.
- As $\mathbf{y} = \mathbf{A}\mathbf{x}$, this yields an eigenvalue problem: $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.
- The characteristic equation

$$(3) \qquad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 5-\lambda & 3 \\ 3 & 5-\lambda \end{vmatrix} = (5-\lambda)^2 - 9 = 0.$$

- Eigenvalues: $\lambda_1 = 8$ and $\lambda_2 = 2$.
- Eigenvectors (which are also the two principle axes): $\underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\text{For } \lambda_1 = 8}$ and $\underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{\text{For } \lambda_2 = 2}$

(Eigenvalues provides factor of stretching along the principle axis)

- Resulting shape after stretching is an ellipse.

RECALL: EXAMPLE 13 in Section 7.2

Markov Process. Powers of a Matrix. Stochastic Matrix

Suppose that the 2004 state of land use in a city of 60 mi^2 of built-up area is

C: Commercially Used 25%   I: Industrially Used 20%   R: Residentially Used 55%.

Find the states in 2009, 2014, and 2019, assuming that the transition probabilities for 5-year intervals are given by the matrix $\mathbf{A}$ and remain practically the same over the time considered.

Five year land use transition matrix

$$\begin{array}{cccc} & \text{From C} & \text{From I} & \text{From R} & \\ \mathbf{A} = & \left[\begin{matrix} 0.7 & 0.1 & 0 \\ 0.2 & 0.9 & 0.2 \\ 0.1 & 0 & 0.8 \end{matrix}\right] & & & \begin{matrix} \text{To C} \\ \text{To I} \\ \text{To R} \end{matrix} \end{array}$$

- Land usage vector (in %) in 2004:

$$\text{LUV}_{2004} = \begin{bmatrix} 25 \\ 20 \\ 55 \end{bmatrix}$$

- Land usage vector (in %) in 2009:

$$\text{LUV}_{2009} = \text{A} \ \ \text{LUV}_{2004} = \begin{bmatrix} 19.5 \\ 34 \\ 46.5 \end{bmatrix}$$

- Land usage vector (in %) in 2014:

$$LUV_{2014} = A\ \ LUV_{2009} = A^2\ \ LUV_{2004} = \begin{bmatrix} 17.05 \\ 43.8 \\ 39.15 \end{bmatrix}$$

- Land usage vector (in %) in 2019:

$$LUV_{2019} = A\ \ LUV_{2014} = A^3\ \ LUV_{2004} = \begin{bmatrix} 16.315 \\ 50.66 \\ 33.025 \end{bmatrix}$$

- Question: will Land usage vector (in %) converge to some fixed vector for large N? That is

$$\mathbf{A}^N\ \ LUV_{2004} = LUV_{SS}$$

Steady state

or

$$\underbrace{\mathbf{A}\ LUV_{SS} = LUV_{SS}}$$

LUV_{SS} not changed by further multiplication of **A**

- To solve this eigenvalue problem, we need to ask:
  - Is 1 indeed an eigenvalue of **A**?
  - If so, LUV_{SS} is corresponding eigenvector of **A**!

Solution:

- Does matrix **A** has an eigenvalue of value 1?
  - Characteristic polynomial:

$$\det(\mathbf{A}-\lambda\mathbf{I}) = \begin{vmatrix} 0.7-\lambda & 0.1 & 0 \\ 0.2 & 0.9-\lambda & 0.2 \\ 0.1 & 0 & 0.8-\lambda \end{vmatrix}$$

$$= 0.49 - 1.89\lambda + 2.4\lambda^2 - \lambda^3 = 0$$

→ $\lambda = 1$ is indeed eigenvalue of **A**

- Eigenvector corresponding to eigenvalue $\lambda = 1$?
  - Substitute $\lambda = 1$ into This gives the value of LUV_{SS}.

$$(\mathbf{A}-\lambda\mathbf{I})\mathbf{x} = \begin{bmatrix} -0.3 & 0.1 & 0 \\ 0.2 & -0.1 & 0.2 \\ 0.1 & 0 & -0.2 \end{bmatrix}\mathbf{x} = \mathbf{0} \xrightarrow[\text{row operations}]{\text{Elementary}} \begin{bmatrix} -0.3 & 0.1 & 0 \\ 0 & -0.1 & 0.6 \\ 0 & 0 & 0 \end{bmatrix}\mathbf{x} = 0$$

$$\Rightarrow \mathbf{x} = c\begin{bmatrix} 2 \\ 6 \\ 1 \end{bmatrix}$$

Value of c to choose for LUV_{SS}?

8.3 Symmetric, Skew-Symmetric, and Orthogonal Matrices

- Eigenvalues and hence eigenvector of a general matrix may be real or complex
- Example: Given $\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 0 & -3 \\ -1 & 1 & 4 \end{bmatrix}$

Eigenvalues and eigenvectors?

<u>Solution:</u>

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

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

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

Eigenvalues: $\lambda_1 = 3, \lambda_2 = 1 + j\sqrt{2}, \lambda_3 = 1 - j\sqrt{2}$

Eigenvectors:

For $\lambda_1 = 3, \qquad \mathbf{e}_1 = k_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

For

$$\lambda_2 = 1 + j\sqrt{2}, \qquad \mathbf{e}_2 = k_2 \begin{bmatrix} -j\sqrt{2} \\ -3 \\ 1 \end{bmatrix}$$

For

$$\lambda_3 = 1 - j\sqrt{2}, \qquad \mathbf{e}_2 = k_2 \begin{bmatrix} +j\sqrt{2} \\ -3 \\ 1 \end{bmatrix}$$

- For a ***square*** general matrix with ***real*** entries, its eigenvalues and hence eigenvector may be real or complex
- Question: are there more specificity on eigenvalues of **special** matrices?

Definitions

Symmetric, Skew-Symmetric, and Orthogonal Matrices

A ***real*** square matrix $\mathbf{A} = [a_{jk}]$ is called

- **Symmetric** if transposition leaves it unchanged,

(1) $\mathbf{A}^\mathsf{T} = \mathbf{A}$, thus $a_{kj} = a_{jk}$,

- **Skew-symmetric** if transposition gives the negative of **A**,

(2) $\mathbf{A}^\mathsf{T} = -\mathbf{A}$, thus $a_{kj} = -a_{jk}$,

- **Orthogonal** if transposition gives the inverse of **A**,

(3) $\mathbf{A}^\mathsf{T} = \mathbf{A}^{-1}$.

- Examples: Symmetric Skew-symmetric Orthogonal

$$\begin{bmatrix} -3 & 1 & 5 \\ 1 & 0 & -2 \\ 5 & -2 & 4 \end{bmatrix}, \quad \begin{bmatrix} 0 & 9 & -12 \\ -9 & 0 & 20 \\ 12 & -20 & 0 \end{bmatrix}, \quad \begin{bmatrix} \tfrac{2}{3} & \tfrac{1}{3} & \tfrac{2}{3} \\ -\tfrac{2}{3} & \tfrac{2}{3} & \tfrac{1}{3} \\ \tfrac{1}{3} & \tfrac{2}{3} & -\tfrac{2}{3} \end{bmatrix}$$

- Note: Any real square matrix **A** may be written as the sum of a symmetric matrix **R** and a skew-symmetric matrix **S**, i.e.,

$$A = R + S$$

where

$$\mathbf{R} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^{\mathsf{T}}) \qquad \text{and} \qquad \mathbf{S} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^{\mathsf{T}}). \tag{4}$$

Example

$$\mathbf{A} = \begin{bmatrix} 9 & 5 & 2 \\ 2 & 3 & -8 \\ 5 & 4 & 3 \end{bmatrix} = \mathbf{R} + \mathbf{S} = \begin{bmatrix} 9.0 & 3.5 & 3.5 \\ 3.5 & 3.0 & -2.0 \\ 3.5 & -2.0 & 3.0 \end{bmatrix} + \begin{bmatrix} 0 & 1.5 & -1.5 \\ -1.5 & 0 & -6.0 \\ 1.5 & 6.0 & 0 \end{bmatrix}$$

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.

Meaning of Theorem 1:

- A general real-valued square matrix may give rise to both real and complex eigenvalues and eigenvectors
- For symmetric and skew-symmetric matrix, however, we know something more about their eigenvalues

(a) The eigenvalues of a symmetric matrix are real. Why?

Assuming symmetric matrix **A** having complex eigenvalue $\lambda^{(c)}$ and eigenvector $\mathbf{e}^{(c)}$, then

$$\mathbf{A}\mathbf{e}^{(c)} = \lambda^{(c)}\mathbf{e}^{(c)}$$

Take complex conjugate

$$\mathbf{A}\overline{\mathbf{e}}^{(c)} = \overline{\lambda}^{(c)}\overline{\mathbf{e}}^{(c)}$$

Multiple on left by $\left(\mathbf{e}^{(c)}\right)^{\mathrm{T}}$

$$\mathbf{e}^{(c)\mathrm{T}}\mathbf{A}\overline{\mathbf{e}}^{(c)} = \mathbf{e}^{(c)\mathrm{T}}\overline{\lambda}^{(c)}\overline{\mathbf{e}}^{(c)}$$

$$= \overline{\lambda}^{(c)}\left(\mathbf{e}^{(c)\mathrm{T}}\overline{\mathbf{e}}^{(c)}\right)$$

$\overline{\lambda}^{(c)}$ complex conjugate of $\lambda^{(c)}$

$\overline{\mathbf{e}}^{(c)}$ complex conjugate of $\mathbf{e}^{(c)}$

Take Transpose

$$\mathbf{e}^{(c)\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \lambda^{(c)}\mathbf{e}^{(c)\mathrm{T}}$$

Matrix **A** symmetric

$$\mathbf{e}^{(c)\mathrm{T}}\mathbf{A} = \lambda^{(c)}\mathbf{e}^{(c)\mathrm{T}}$$

Multiple on right by $\overline{\mathbf{e}}^{(c)}$

$$\mathbf{e}^{(c)\mathrm{T}}\mathbf{A}\overline{\mathbf{e}}^{(c)} = \lambda^{(c)}\left(\mathbf{e}^{(c)\mathrm{T}}\overline{\mathbf{e}}^{(c)}\right)$$

$$\mathbf{e}^{(c)\mathrm{T}}\mathbf{A}\overline{\mathbf{e}}^{(c)} = \overline{\lambda}^{(c)}\left(\mathbf{e}^{(c)\mathrm{T}}\overline{\mathbf{e}}^{(c)}\right) = \lambda^{(c)}\left(\mathbf{e}^{(c)\mathrm{T}}\overline{\mathbf{e}}^{(c)}\right)$$

$\Rightarrow$ $\overline{\lambda}^{(c)} = \lambda^{(c)}$ must be real!

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

Assume skew-symmetric matrix **A** having complex eigenvalue $\lambda^{(c)}$ and hence eigenvector $\mathbf{e}^{(c)}$, then

$$\mathbf{A}\mathbf{e}^{(c)} = \lambda^{(c)}\mathbf{e}^{(c)}$$

Take complex conjugate

$$\mathbf{A}\overline{\mathbf{e}}^{(c)} = \overline{\lambda}^{(c)}\overline{\mathbf{e}}^{(c)}$$

Multiple on left by $\left(\mathbf{e}^{(c)}\right)^{\mathrm{T}}$

$$\mathbf{e}^{(c)\mathrm{T}}\mathbf{A}\overline{\mathbf{e}}^{(c)} = \mathbf{e}^{(c)\mathrm{T}}\overline{\lambda}^{(c)}\overline{\mathbf{e}}^{(c)} = \overline{\lambda}^{(c)}\left(\mathbf{e}^{(c)\mathrm{T}}\overline{\mathbf{e}}^{(c)}\right)$$

$\overline{\lambda}^{(c)}$ complex conjugate of $\lambda^{(c)}$

$\overline{\mathbf{e}}^{(c)}$ complex conjugate of $\mathbf{e}^{(c)}$

Take Transpose

$$\mathbf{e}^{(c)\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \lambda^{(c)}\mathbf{e}^{(c)\mathrm{T}}$$

Matrix **A** skew-symmetric

$$-\mathbf{e}^{(c)\mathrm{T}}\mathbf{A} = \lambda^{(c)}\mathbf{e}^{(c)\mathrm{T}}$$

Multiple on right by $\overline{\mathbf{e}}^{(c)}$

$$\mathbf{e}^{(c)\mathrm{T}}\mathbf{A}\overline{\mathbf{e}}^{(c)} = -\lambda^{(c)}\left(\mathbf{e}^{(c)\mathrm{T}}\overline{\mathbf{e}}^{(c)}\right)$$

$$\mathbf{e}^{(c)\mathrm{T}}\mathbf{A}\overline{\mathbf{e}}^{(c)} = \overline{\lambda}^{(c)}\left(\mathbf{e}^{(c)\mathrm{T}}\overline{\mathbf{e}}^{(c)}\right) = -\lambda^{(c)}\left(\mathbf{e}^{(c)\mathrm{T}}\overline{\mathbf{e}}^{(c)}\right)$$

$\Rightarrow$ $\overline{\lambda}^{(c)} = -\lambda^{(c)}$ must be pure imaginary complex number or zero!

- Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 7 & 3 \\ 7 & 4 & -5 \\ 3 & -5 & 6 \end{bmatrix}$$

Eigenvalues:
-7.0079, 7.0359, 10.9720

- Example:

$$\mathbf{A} = \begin{bmatrix} 0 & 7 & 3 \\ 7 & 0 & -5 \\ 3 & -5 & 0 \end{bmatrix}$$

Eigenvalues:
-10.1798, 2.7925, 7.3874

- Example:

$$\mathbf{A} = \begin{bmatrix} 0 & 7 & 3 \\ -7 & 0 & -5 \\ -3 & 5 & 0 \end{bmatrix}$$

Eigenvalues:
0, 9.1104i, - 9.1104i

Orthogonal Transformations and Orthogonal Matrices

- **Orthogonal transformations** are transformations

(5) $$\mathbf{y} = \mathbf{A}\mathbf{x}$$

where **A** is an (*square*) orthogonal matrix.

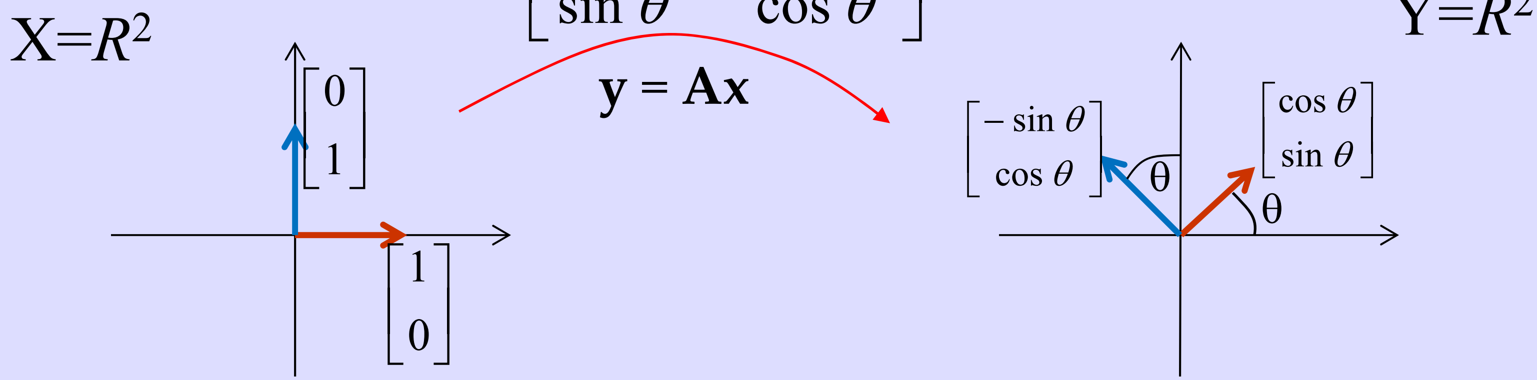
- Geometrically, orthogonal transformation maps each vector **x** in R^n to a vector **y** in R^n through the rotation of an angle.

- Example: the following mapping from **x** in R^2 to **y** in R^2

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

is an orthogonal transformation of rotating an angle of θ.

$$\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$



- **A** is a counter-clockwise rotation of angle θ
- $\mathbf{A}^{\mathrm{T}}$ is a clockwise rotation of angle θ, because

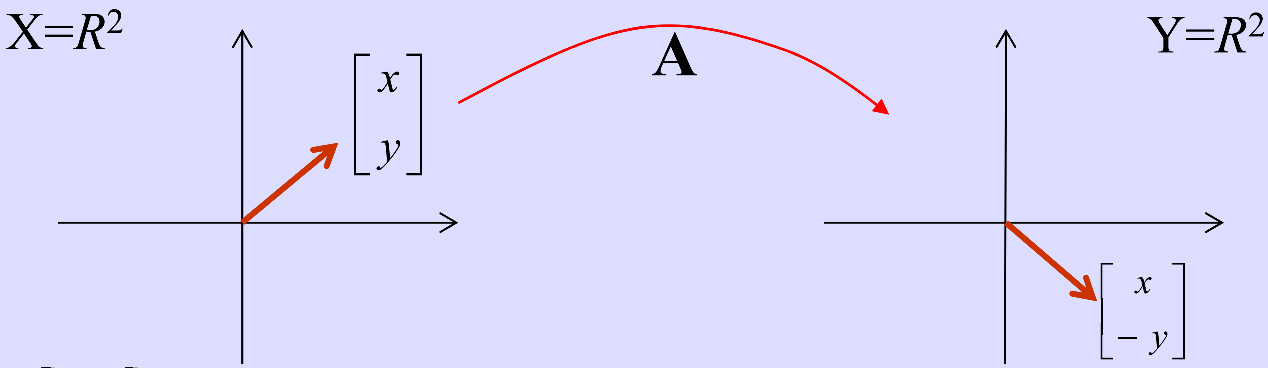
$$\mathbf{A}^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}$$

- Hence, $\mathbf{A}\mathbf{A}^{\mathrm{T}}=\mathbf{A}^{\mathrm{T}}\mathbf{A}=\mathbf{I}$
- Orthogonal transformation in R^2 or in R^3 is a **rotation** (possibly combined with a reflection about a straight line or a plane).

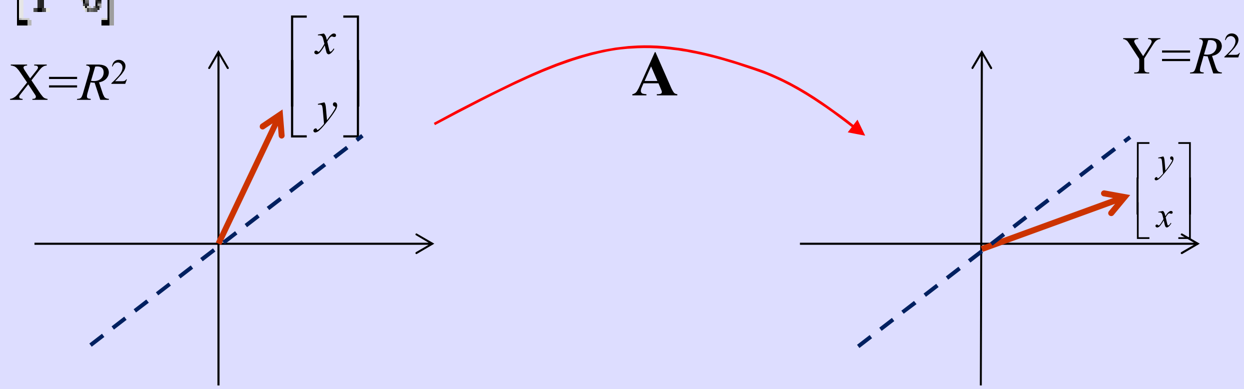
- Other examples of orthogonal transformation

$$\mathbf{A}=\begin{bmatrix} 0 & -0.80 & -0.60 \\ 0.80 & -0.36 & 0.48 \\ 0.60 & 0.48 & -0.64 \end{bmatrix} \qquad \begin{pmatrix} \text{rotoinversion:} \\ \text{axis } (0, -3/5, 4/5), \text{ angle } 90^\circ \end{pmatrix}$$

$$\mathbf{A}=\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \text{(reflection across } x\text{-axis)}$$



$$\mathbf{A}=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \qquad \text{(reflection about the line } 45^{\circ}\text{)}$$

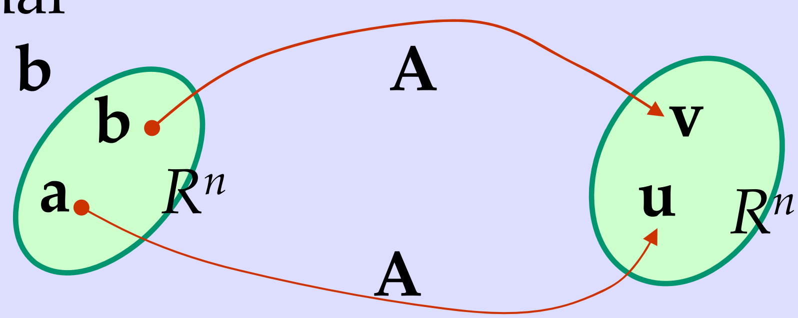


Properties of Orthogonal Transformation

Theorem 2 Invariance of Inner Product

An orthogonal transformation preserves the value of the **inner product** of vectors **a** and **b** in R^n.

- With Inner product defined in R^n as: $(\mathbf{p} \bullet \mathbf{q}) = \mathbf{p}^T \mathbf{q}$
- For **a** and **b** in R^n, orthogonal $n \times n$ matrix **A** maps **a** and **b** into $\mathbf{u} = \mathbf{Aa}$ and $\mathbf{v} = \mathbf{Ab}$, then $\mathbf{u} \cdot \mathbf{v} = \mathbf{a} \cdot \mathbf{b}$.



- For $\mathbf{u} = \mathbf{Aa}$, $\|\mathbf{u}\| = \sqrt{(\mathbf{Aa} \bullet \mathbf{Aa})} = \sqrt{(\mathbf{a}^T \mathbf{A}^T \mathbf{Aa})} =.\sqrt{\mathbf{a} \bullet \mathbf{a}} = \|\mathbf{a}\|$
  → orthogonal transformation also preserves the **length** or **norm** of vectors mapping to one another in R^n

Theorem 3 Orthonormality of Column and Row Vectors

A real square matrix is orthogonal if and only if its column vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ (and also its row vectors) form an **orthonormal system**, that is,

$$(10) \qquad \mathbf{a}_j \bullet \mathbf{a}_k = \mathbf{a}_j^T \mathbf{a}_k = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases}$$

Proof:

$$\underbrace{\begin{bmatrix} \leftarrow & \mathbf{a}_1^T & \rightarrow \\ \leftarrow & \mathbf{a}_2^T & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{a}_n^T & \rightarrow \end{bmatrix}}_{\mathbf{A^T=A^{-1}}} \underbrace{\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}}_{\mathbf{A}} == \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Theorem 4

Determinant of an Orthogonal Matrix

The determinant of an orthogonal matrix has the value +1 or −1.

Proof: With $\det(\mathbf{A}^T)\det(\mathbf{A})=\det(\mathbf{A}^T\mathbf{A})=\det(\mathbf{I})=1$

and $\det(\mathbf{A}^T)=\det(\mathbf{A})$

we have

$$\det(\mathbf{A})\det(\mathbf{A}) = 1 \rightarrow \det(\mathbf{A}) = +1 \text{ or } -1$$

Theorem 5 Eigenvalues of an Orthogonal Matrix

The eigenvalues of an orthogonal matrix **A** have absolute value 1 (they may be real or complex as for any general matrix).

Proof:

Consider an eigenvlaue λ and its eigenvector **e** of an orthogonal matrix **A**, i.e.,

$$\mathbf{A}\mathbf{e} = \lambda\,\mathbf{e}\,.$$

Let $\bar{\lambda}$ and $\bar{\mathbf{e}}$ be complex conjugate of λ and **e**. We have

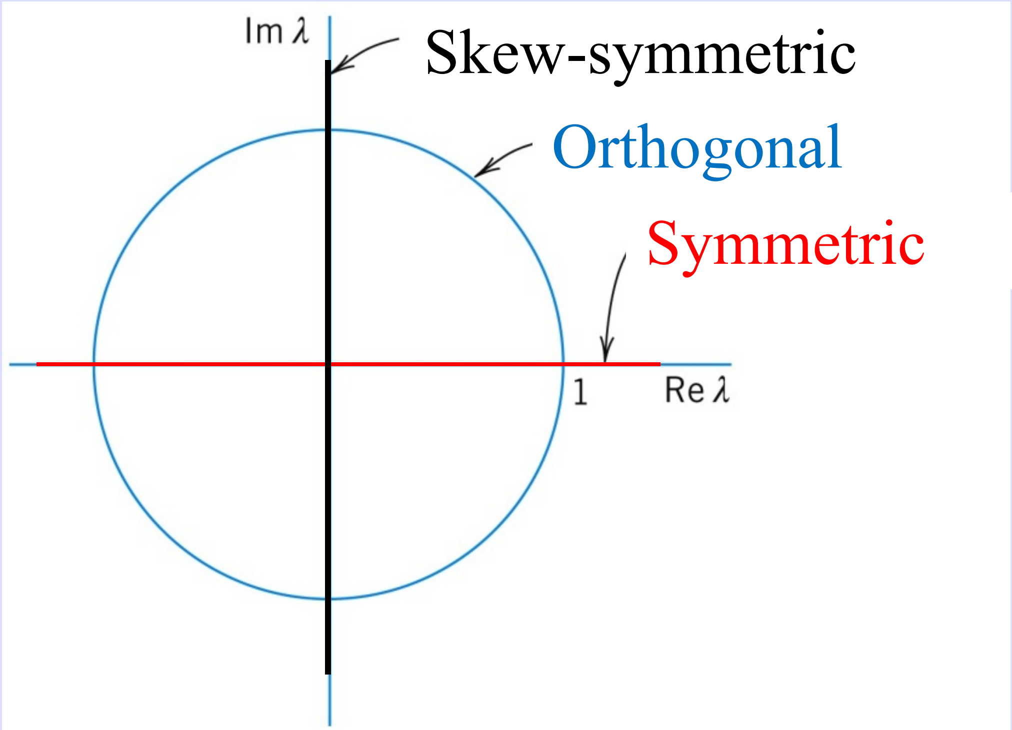
$$\mathbf{A}\bar{\mathbf{e}} = \bar{\lambda}\,\bar{\mathbf{e}}\,, \quad \text{or} \quad (\mathbf{A}\bar{\mathbf{e}})^{\mathrm{T}} = (\bar{\lambda}\,\bar{\mathbf{e}})^{\mathrm{T}}$$

Hence, $(\mathbf{A}\bar{\mathbf{e}})^{\mathrm{T}}\mathbf{A}\mathbf{e} = (\bar{\lambda}\,\bar{\mathbf{e}})^{\mathrm{T}}(\lambda\mathbf{e})$

$$\Rightarrow \bar{\mathbf{e}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{e} = |\lambda|^2(\bar{\mathbf{e}}^{\mathrm{T}}\mathbf{e})$$

$$\Rightarrow \quad |\mathbf{e}|^2 = |\lambda|^2|\mathbf{e}|^2 \Rightarrow |\lambda|^2 = 1 \Rightarrow |\lambda| = 1$$

Summary



Characterization of eigenvalues of Symmetric, skew-symmetric and orthogonal matrix

8.4 Eigenbases. Diagonalization. Quadratic Forms

Recall: Basis of Vector Space (Section 7.4)

- Given a vector space V with **dim** $V = N$, any vector set containing N **linearly independent** vectors in V forms a **basis** for V.
- For example: V= R^3 is a vector space of Dim V=3. The following vector sets are basis:

Proof of LI?

  - S= ($\mathbf{a}_{(1)=}$[1, 0, 1], $\mathbf{a}_{(2)=}$[2, 0, 1], $\mathbf{a}_{(3)=}$[1, 1, 2])
  - S= ($\mathbf{a}_{(1)=}$[10, 4, 0], $\mathbf{a}_{(2)=}$[0, 5, 1], $\mathbf{a}_{(3)=}$[1, 1, 5])
- Given a basis of V, then any vector **a** in V can be (uniquely) represented as a linear combination of the basis vectors $\mathbf{a}_{(i)}$, i=1, …, N, i.e.,

$$\mathbf{a} = c_1\mathbf{a}_{(1)} + c_2\mathbf{a}_{(2)} + \ldots.. + c_N\mathbf{a}_{(N)}$$

- Especially, an **orthonormal basis** is a basis with unit length basis vectors that are orthogonal to each other:

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

- Example: S= ($\mathbf{a}_{(1)}$=[1, 0, 0], $\mathbf{a}_{(2)}$=[0, 1, 0], $\mathbf{a}_{(3)}$=[0, 0, 1]) is an orthonormal basis for V= R^3.

- **Advantage of orthonormal basis**: On representing any vector **a** as linear combination of a basis ($\mathbf{a}_{(i)}$, i=1, …, N):

$$\mathbf{a} = c_1\mathbf{a}_{(1)} + c_2\mathbf{a}_{(2)} + \ldots.. + c_N\mathbf{a}_{(N)}\,,$$

  If the basis is orthonormal, the constants c_i, i=1,…, N, can be simply obtained as

$$c_i = \mathbf{a}_{(i)} \bullet \mathbf{a} = \mathbf{a}_{(i)}^{\mathrm{T}} \mathbf{a}$$

  Otherwise, constants c_i, i=1, …, N, may need to be determined by solving a set of linear equations.

Question: Given an $n \times n$ matrix **A,** will the eigenvectors of matrix **A** form an basis for R^n ?

Answer:

- Generally, eigenvectors of an general $n \times n$ matrix **A** do not form a basis for R^n.
- In many cases, eigenvectors of an $n \times n$ matrix **A** do form a basis for R^n. Such basis formed is called "**eigenbasis**" (basis of eigenvectors) of R^n.
- Specifically, we have the following facts:
  - When a matrix **A** has distinct eigenvalues, its eigenvectors will definitely form eigenbasis
  - When a matrix **A** has multiple eigenvalues, it eigenvectors may or may not form eigenbasis
  - Symmetric matrices **always** yield eigenbasis regardless of having distinct/multiple eigenvalues

Theorem 1

Basis of Eigenvectors

If an general $n \times n$ matrix $\mathbf{A}$ has n ***distinct*** eigenvalues, then eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ of $\mathbf{A}$ form a basis (hence **eigenbasis**) for R^n.

Example: 2 × 2 Matrix with Distinct Roots

- Given the 2 × 2 matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix}.$$

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

- Eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -3$. Eigenvalues are distinct → by Theorem 1, eigenvectors form **eigenbasis** for R^2.

- Checking:

$$\text{For} \quad \lambda_1 = -1, \quad \mathbf{e}_1 = k\begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{for any } k \neq 0$$

$$\text{For} \quad \lambda_2 = -3, \quad \mathbf{e}_2 = k\begin{bmatrix} 1 \\ -3 \end{bmatrix} \text{for any } k \neq 0$$

$\mathbf{e}_1$ and $\mathbf{e}_2$ linearly independent and form an **eigenbasis** for R^2.

Example: 3 × 3 Matrix with Distinct Roots

- Given the 3×3 matrix

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

$$\det(\mathbf{A} - \lambda\mathbf{I}) == \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 3-\lambda & 2 \\ -1 & 1 & 2-\lambda \end{vmatrix} = -\lambda^3 + 7\lambda^2 - 14\lambda + 8 = 0$$

- Eigenvalues λ_1=1, λ_2= 2 and λ_3=4 → Eigenvalues distinct
  → by Theorem 1, eigenvectors form **eigenbasis** for R^2.
- Checking:

$$\text{For } \lambda_1 = 1, \quad \mathbf{e}_1 = k \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \text{ for any } k \neq 0$$

- Checking:

$$\text{For} \quad \lambda_2 = 2, \quad \mathbf{e}_2 = k\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \text{ for any } k \neq 0$$

$$\text{For} \quad \lambda_3 = 4, \quad \mathbf{e}_3 = k\begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} \text{ for any } k \neq 0$$

$\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ linearly independent and form an **eigenbasis** for R^3.

Theorem 1 is for matrices with distinct eigenvalues, how about matrices with multiple eigenvalues?

EXAMPLE 2 of Section 8.1 revisited

3 × 3 Matrix with Multiple Roots

- Given the matrix:

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

- Characteristic equation

Multiplicity=2 for this eigenvalue

$$-\lambda^3 - \lambda^2 + 21\lambda + 45 = 0.$$

yields the eigenvalues of **A**) $\lambda_1 = 5, \lambda_2 = \lambda_3 = -3.$

- Multiple root case. Do the eigenvectors form an **eigenbasis** for R^3? In this case, yes.

- Checking:

  For $\lambda_1 = 5$, $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$

  For $\lambda_{2,3} = -3$, $\mathbf{e}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_3 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$

  With $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ LI → they form a basis for R^3.

- In this case, the eigenvalue $\lambda = -3$ with multiplicity=2 yields 2 linearly independent eigenvectors $\mathbf{e}_2$ and $\mathbf{e}_3$.
- Hence, we have a total of 3 linearly independent eigenvectors (1 for $\lambda = 5$, and 2 for $\lambda = -3$), and the formation of an **eigenbasis** for R^3.

- Generally, however, there are matrices with eigenvalue of multiplicity p that does not yield p linearly independent eigenvectors.
- These matrices do not have enough linearly independent eigenvectors to form an **eigenbasis**.

Example: Matrix with Multiple Roots, Again

- Consider the matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

- Characteristic equation

$$\det(\mathbf{A} - \lambda \mathbf{I}_2) = \det \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$$

- The eigenvalues of **A**: $\lambda_1 = \lambda_2 = 1.$ Multiplicity of 2
- Do the eigenvectors form an **eigenbasis** for R^2?

- Eigenvectors: For $\lambda_1 = \lambda_2 = 1$.

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{e} = \mathbf{0}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\mathbf{e} = \mathbf{0}, \quad \mathbf{e} = k\begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ for any } k \neq 0$$

- In this case, there is one eigenvector → No **eigenbasis**!
- <u>One may wonder</u>: Is there a second eigenvector of **A**?
  <u>Answer:</u> The concept of ***Generalized Eigenvector***.
- Consider vector $\mathbf{v} = k\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ orthogonal to **e**, we have $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{e}$.
- Hence, **v** is such that $(\mathbf{A} - \lambda\mathbf{I})^2\mathbf{v} = \mathbf{0}$. **v** is called the ***generalized eigenvector*** of order 2.
- This compared to **e**, which satisfies $(\mathbf{A} - \lambda\mathbf{I})\mathbf{e} = \mathbf{0}$ and is a ***generalized vector*** of order 1.

- Hence, for general $n \times n$ matrix **A**, **eigenbasis** is guaranteed only for those matrices with ***distinct*** eigenvalues only.
- For general $n \times n$ matrix **A**, **eigenbasis** is **not** guaranteed for those matrices with ***multiple*** eigenvalues.
- However, if matrix **A** is **symmetric**,
  - Its eigenvectors will **always give rise to an eigenbasis** for R^n, regardless of whether matrix **A** has distinct or multiple eigenvalues.
  - Its eigenvectors will more readily allow **orthonormal eigenbasis** to be generated.

Nice Properties of Eigenvectors of Symmetric Matrix in case of distinct eigenvalues

For symmetric matrix **A**, eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof: Let $\mathbf{e}_j$ and $\mathbf{e}_k$ be eigenvectors of **A** corresponding to distinct eigenvalues λ_j and λ_k, $k \neq j$, then

$$\mathbf{e}_k{}^{\mathrm{T}}\mathbf{A}\mathbf{e}_j = \mathbf{e}_k{}^{\mathrm{T}}\left(\mathbf{A}\mathbf{e}_j\right) = \mathbf{e}_k{}^{\mathrm{T}}\left(\lambda_j \mathbf{e}_j\right) = \lambda_j \mathbf{e}_k{}^{\mathrm{T}}\mathbf{e}_j$$

$$\mathbf{e}_k{}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{e}_j = \left(\mathbf{A}\mathbf{e}_k\right)^{\mathrm{T}}\mathbf{e}_j = \left(\lambda_k \mathbf{e}_k\right)^{\mathrm{T}}\mathbf{e}_j = \lambda_k \mathbf{e}_k{}^{T}\mathbf{e}_j$$

Since $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$, the LHS of the above two equations are equal, subtracting the two equation yields

$$0 = \left(\lambda_j - \lambda_k\right)\mathbf{e}_k{}^{T}\mathbf{e}_j$$

Since the eigenvalues are distinct, $\lambda_j \neq \lambda_k$, this implies $\mathbf{e}_k{}^{\mathrm{T}}\mathbf{e}_j = 0 \rightarrow$ Eigenvector $\mathbf{e}_k$ and $\mathbf{e}_j$ are orthogonal for $k \neq j$.

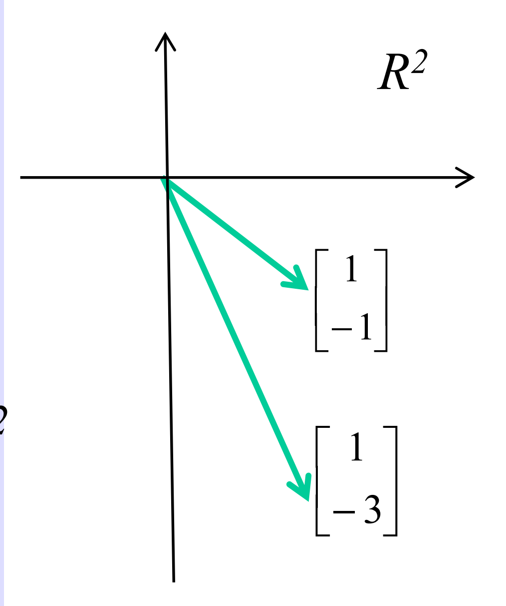
Example: Non-Symmetric Matrix with Distinct Roots

- Consider the matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix}.$
- Eigenvalues: $\lambda_1 = -1, \lambda_2 = -3$ → eigenvalues distinct
  → By Theorem 1, eigenvectors form **eigenbasis** for R^2.
- Eigenvectors

  For $\lambda_1 = -1, \quad \mathbf{e}_1 = k\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ for any $k \neq 0$

  For $\lambda_2 = -3, \quad \mathbf{e}_2 = k\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ for any $k \neq 0$

  $\mathbf{e}_1$ and $\mathbf{e}_2$ are LI and form a basis for R^2 but they are not orthogonal.



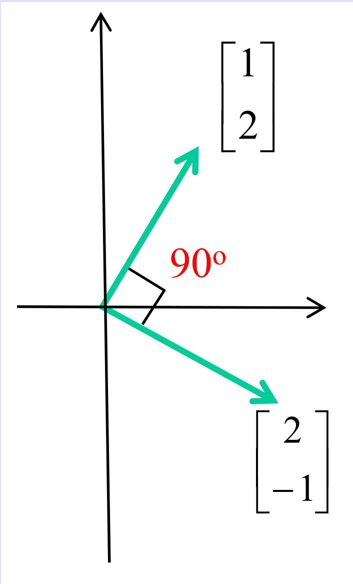
Example: Symmetric matrix with Distinct Roots

- Given the symmetric matrix $\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 2 & 7 \end{bmatrix}$,
- Eigenvalues: $\lambda_1=3$ and $\lambda_2=8$ → distinct eigenvalues!
- Eigenvectors: $\mathbf{e}_1 = k_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and $\mathbf{e}_2 = k_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

  $\mathbf{e}_1$ and $\mathbf{e}_2$ form an eigenbasis for R^2.

- Note: in this symmetric $\mathbf{A}$ case, $\mathbf{e}_1$ and $\mathbf{e}_2$ are not only LI, they are also orthogonal → one can simply form an orthonormal basis by dividing them with their length:

$$\hat{\mathrm{e}}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \hat{\mathrm{e}}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$



Another Example: Symmetric Matrix with Distinct eigenvalues – Revisit Example 1 of Section 8.2

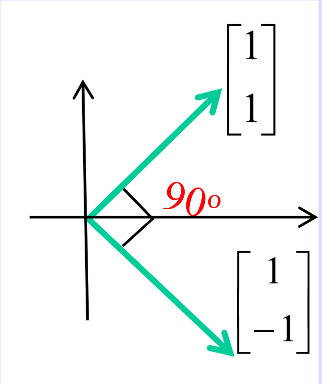
- The matrix (of elastic membrane extension)

$$\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \Rightarrow \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 5-\lambda & 3 \\ 3 & 5-\lambda \end{vmatrix} = (\lambda - 8)(\lambda - 2)$$

- Eigenvectors:

$$\text{For } \lambda_1 = 8,\ \mathbf{e}_1 = k\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ for any } k \neq 0$$

$$\text{For } \lambda_2 = 2,\ \mathbf{e}_2 = k\begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ for any } k \neq 0$$



- Again, $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal (not only LI) and readily form an orthonormal eigenbasis via division by their length:

$$\hat{\mathbf{e}}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix},\ \hat{\mathbf{e}}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Advantage of Symmetric Over Non-Symmetric Matrix

- For general matrix **A**, eigenvectors corresponding to distinct eigenvalues are LI
  → they form eigenbasis
  → they can be used to form orthonormal eigenbasis (by going through some needed computation)

- For symmetric matrix **A**, eigenvectors corresponding to distinct eigenvalues are **already** orthogonal
  → they form orthogonal eigenbasis
  → they readily yield orthonormal eigenbasis

Nice Properties of Eigenvectors of Symmetric Matrix
in the case of multiple eigenvalues

If an eigenvalue λ of a symmetric matrix $\mathbf{A}$ is of multiplicity k, then

- The eigenspace corresponding to λ must be of dimension k, i.e., there must be k **linearly independent** eigenvectors corresponding to λ.

  (Situation such as Example $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ won't occur in symmetric matrix.)

- One can always use these eigenvectors to construct an orthonormal basis for this k-dimension eigenspace of λ.
- Together with the eigenvectors of the distinct eigenvalues yields orthonormal eigenbasis for R^n.

Example: Symmetric Matrix with Multiple eigenvalues

- Find the eigenbasis of the matrix

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

- Solution:
  - Characteristic equation:

$$\det(\mathbf{A} - \lambda \mathrm{I}) = 0 \ \rightarrow \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = 0$$

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

$$\rightarrow \lambda^3 - 6\lambda^2 + 9\lambda - 4 = 0$$

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

$$\rightarrow \text{eigenvalues}: \lambda_1 = 4; \lambda_2 = 1, 1 \quad \text{(Multiplicity=2)}$$

- Eigenvector corresponding to $\lambda_1 = 4$:

$$(\mathbf{A}-4\mathbf{I})\mathbf{e}_1 = \mathbf{0} \Rightarrow \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}\mathbf{e}_1 = \mathbf{0} \Rightarrow \mathbf{e}_1 = k\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{for any } k \neq 0$$

- Eigenvector corresponding to $\lambda_2 = 1$ (multiplicity=2):

$$\left.\begin{aligned} &(\mathbf{A}-\mathbf{I})\mathbf{e} = \mathbf{0} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\mathbf{e} = \mathbf{0} \\ &\Rightarrow \mathbf{e} = k_1\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + k_2\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \end{aligned}\right\}$$

Can form

$$\mathbf{e}_2 = k_1\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = k_2\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

for any nonzero k_1, k_2

- <u>Note:</u> $\mathbf{e}_1$ orthogonal to $\mathbf{e}_2$ and $\mathbf{e}_3$ (eigenvectors of different eigenvalues orthogonal to each other)
- $\mathbf{e}_2$ and $\mathbf{e}_3$ are LI and so eigenspace of λ_2 has dim 2.
- Readily form orthonormal basis for R^3 with ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$).

Formation of orthonormal Eigenbasis

Note, in this case, eigenvector corresponding to λ_1 :

$$\mathbf{e}_1 = k\begin{bmatrix}1\\1\\1\end{bmatrix} \text{ for any } k \neq 0 \text{ is orthogonal to } \mathbf{e}_2 \text{ and } \mathbf{e}_3 \text{ corresponding to } \lambda_2$$

However, in this case, eigenvector $\mathbf{e}_2$ is not orthogonal to $\mathbf{e}_3$ because they are not eigenvectors of distinct eigenvalues, but of the same $\lambda_2 = 1$ (of multiplicity 2)

Generally, eigenvectors (in this case $\mathbf{e}_2$ and $\mathbf{e}_3$) corresponding to the same eigenvalue of multiplicity > 1 are not necessarily orthogonal to one another.

However, we can always construct orthogonal basis of the corresponding eigenspace using these eigenvectors, and the resulting orthogonal basis vectors are still eigenvectors corresponding to the particular eigenvalues.

To form orthonormal basis from $\mathbf{e}_2$ and $\mathbf{e}_3$. Keep $\mathbf{e}_2$ as one of the basis vectors, and use a linear combination of $\mathbf{e}_2$ and $\mathbf{e}_3$ to form

$$\tilde{\mathbf{e}}_3 = k_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ (for some } k_1, k_2) \quad \text{such that} \quad \mathbf{e}_2 \cdot \tilde{\mathbf{e}}_3 = \mathbf{e}_2{}^T \tilde{\mathbf{e}}_3 = 0$$

This gives $2k_1 + k_2 = 0$, or $k_2 = -2k_1$, or

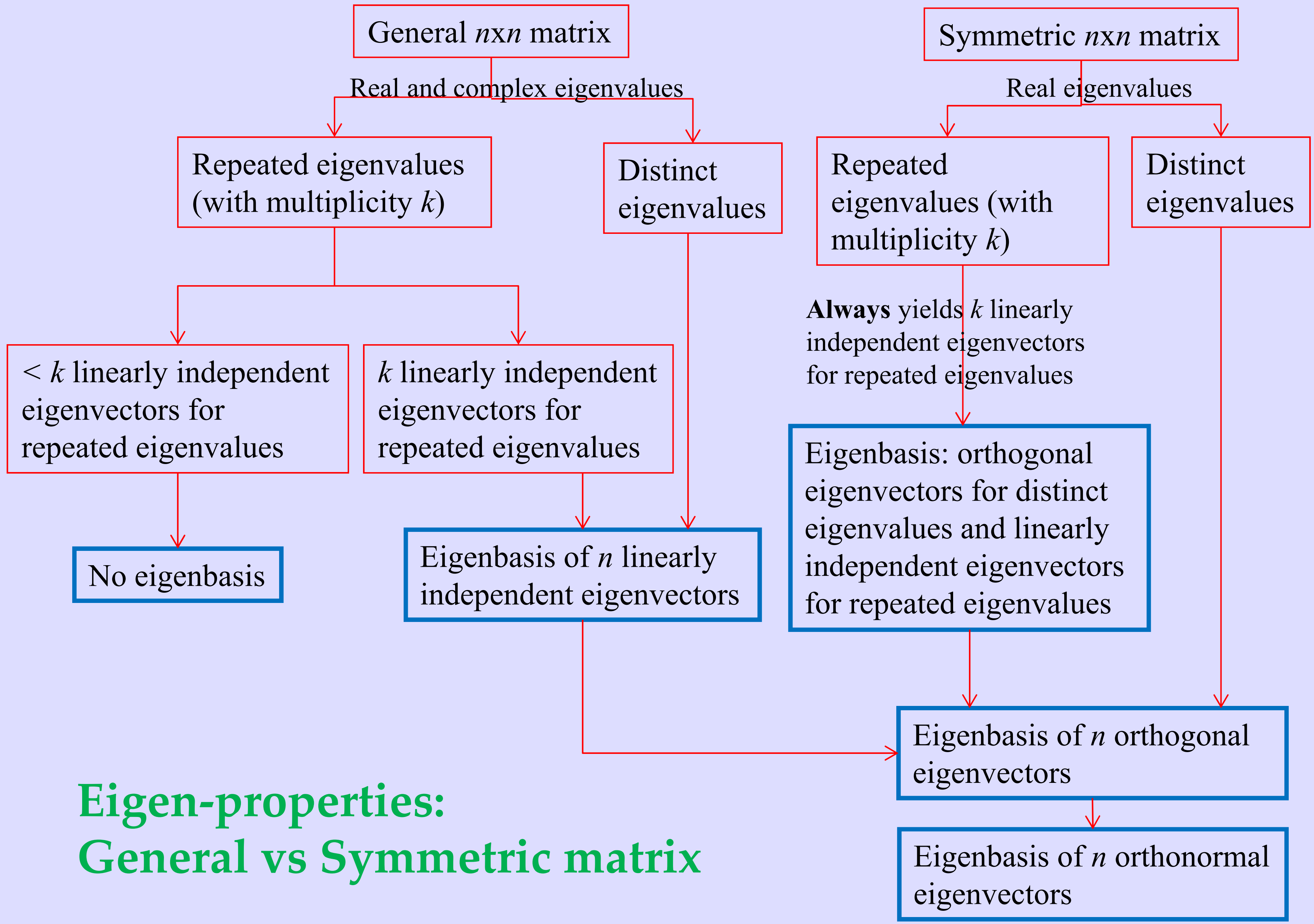
$$\tilde{\mathbf{e}}_3 = k_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} - 2k_1 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = k_1 \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

Then we have the eigenvectors $\{\mathbf{e}_1, \mathbf{e}_2, \tilde{\mathbf{e}}_3\}$ forming an orthognal basis for R^3.

Furthermore, we can form an orthonormal basis making use of the constant k's in $\{\mathbf{e}_1, \mathbf{e}_2, \hat{\mathbf{e}}_3\}$ to make the resulting basis vectors of unit length:

$$\hat{\mathbf{e}}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \hat{\mathbf{e}}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \hat{\mathbf{e}}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

Then $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ forms an orthoronormal eigenbasis for R^3



Eigen-properties:
General vs Symmetric matrix

Similarity of Matrices. Diagonalization

DEFINITION

Similar Matrices. Similarity Transformation

- A general $n \times n$ matrix $\hat{\mathbf{A}}$ is called **similar** to an $n \times n$ matrix $\mathbf{A}$ if

$$\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \tag{4}$$

for some (nonsingular, i.e., invertible) $n \times n$ matrix $\mathbf{P}$.

- The transformation $\mathbf{P}$, which gives $\hat{\mathbf{A}}$ from $\mathbf{A}$, is called a **similarity transformation**.

Theorem 3

Eigenvalues and Eigenvectors of Similar Matrices

- If $\hat{\mathbf{A}}$ is similar to $\mathbf{A}$, then $\hat{\mathbf{A}}$ has the same eigenvalues as $\mathbf{A}$.
- Furthermore, if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, then $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ *i*s an eigenvector of $\hat{\mathbf{A}}$ corresponding to the same eigenvalue λ.

Proof:

- Given that $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ, then

$$\hat{\mathbf{A}}\mathbf{y} = \left(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right)\mathbf{P}^{-1}\mathbf{x} = \mathbf{P}^{-1}\mathbf{A}\mathbf{x} = \mathbf{P}^{-1}\left(\lambda\mathbf{x}\right) = \lambda\left(\mathbf{P}^{-1}\mathbf{x}\right) = \lambda\mathbf{y}$$

Hence, $\mathbf{y}$ is the eigenvector of $\hat{\mathbf{A}}$ corresponding to the same eigenvalue λ.

Theorem 4: Diagonalization of a Matrix

- If the eigenvectors of an general $n \times n$ matrix $\mathbf{A}$ form an eigenbasis, then

$$\text{(5)} \qquad \mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$$

is diagonal, with the eigenvalues of $\mathbf{A}$ as the entries on the main diagonal.

- Here $\mathbf{X}=[\mathbf{e}_1 | \mathbf{e}_2 | \ldots | \mathbf{e}_n]$ is the matrix with eigenvectors $\mathbf{e}_i$, i=1,…, n, of $\mathbf{A}$ as column vectors.

<u>Proof:</u> $\mathbf{e}_i$ is the eigenvector for eigenvalue λ_i , i=1,…, n, hence

$$\begin{aligned}
\text{(5*)} \qquad \mathbf{X}^{-1}\mathbf{A}\mathbf{X} &= \mathbf{X}^{-1}\mathbf{A}[\mathbf{e}_1 | \mathbf{e}_2 | \ldots | \mathbf{e}_n] \\
&= \mathbf{X}^{-1}[\lambda_1\mathbf{e}_1 | \lambda_2\mathbf{e}_2 | \ldots | \lambda_n\mathbf{e}_n] \\
&= \mathbf{X}^{-1} \underbrace{[\mathbf{e}_1 | \mathbf{e}_2 | \ldots | \mathbf{e}_n]}_{\mathbf{X}} \underbrace{\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)}_{\mathbf{D}} = \mathbf{D}
\end{aligned}$$

- Also, we have

$$\textbf{(5*)} \qquad \mathbf{A}^m = \mathbf{X}\mathbf{D}^m\mathbf{X}^{-1} \qquad (m = 2, 3, \dots).$$

<u>Proof:</u> $\mathbf{e}_i$ is the eigenvector for eigenvalue λ_i, i=1,…, n, hence

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$$

$$\rightarrow \mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1}$$

$$\rightarrow \mathbf{A}^m = \underbrace{(\mathbf{X}\mathbf{D}\mathbf{X}^{-1})(\mathbf{X}\mathbf{D}\mathbf{X}^{-1})\dots\dots(\mathbf{X}\mathbf{D}\mathbf{X}^{-1})}_{m}$$

$$\rightarrow \mathbf{A}^m = (\mathbf{X}\underbrace{\mathbf{D}\mathbf{D}\dots\ \mathbf{D}}_{m}\mathbf{X}^{-1})$$

$$\rightarrow \mathbf{A}^m = \mathbf{X}\mathbf{D}^m\mathbf{X}^{-1}$$

- **Note:** Diagonalization may encounter complex eigenvalues and eigenvectors, hence complex **D** and **X**.

EXAMPLE 4 Diagonalization

- Diagonalize

$$\mathbf{A} = \begin{bmatrix} 7.3 & 0.2 & -3.7 \\ -11.5 & 1.0 & 5.5 \\ 17.7 & 1.8 & -9.3 \end{bmatrix}.$$

Solution.

- Characteristic equation: $-\lambda^3 - \lambda^2 + 12\lambda = 0.$
- Eigenvalues: $\lambda_1 = 3, \lambda_2 = -4, \lambda_3 = 0.$
- Eigenvectors: for λ_1 , λ_2 and λ_3

$$\begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix},$$

EXAMPLE 4 (continued)

- Form

$$\mathbf{X} = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}, \qquad \mathbf{X}^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}.$$

- Calculate $\mathbf{X}^{-1}\mathbf{AX}$

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{AX} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix} \begin{bmatrix} -3 & -4 & 0 \\ 9 & 4 & 0 \\ -3 & -12 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- Calculate $\mathbf{A}^4 = \mathbf{XD}^4\mathbf{X}^{-1}$

$$\mathbf{A}^4 = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 3^4 & 0 & 0 \\ 0 & (-4)^4 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix} = \begin{bmatrix} -276.1 & -67.4 & 154.9 \\ 162.7 & 99.8 & -106.3 \\ -941.7 & -169.8 & 513.3 \end{bmatrix}$$

Diagonalization of a Symmetric Matrix

- Diagonalization is more efficient if matrix is symmetric
- In this case, we can more readily determine an orthonormal eigenbasis.
- Hence, $\mathbf{X}$ is formed with the orthonormal eigenvectors of $\mathbf{A}$ as its columns and $\mathbf{X}^{-1}$ is readily obtained as $\mathbf{X}^{-1} = \mathbf{X}^{T}$.

EXAMPLE: Diagonalization of Symmetrix Matrix

- To diagonalize

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & -2 \\ 0 & -2 & 0 \\ -2 & 0 & 3 \end{bmatrix}, \quad \text{with eigenvalues -2, 4, -1}$$

- Eigenvectors (forming orthogonal basis):

$$\mathbf{e}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

- Easily obtain orthonormal eigenbasis for R^3: $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$

$$\hat{\mathbf{e}}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \hat{\mathbf{e}}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \hat{\mathbf{e}}_3 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

EXAMPLE: Diagonalization of Symmetrix Matrix

- Form **X**:

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

(Orthogonal matrix)

Easily obtained
$\mathbf{X}^{-1}$ as $\mathbf{X}^{T}$

- With $\mathbf{X}^{-1} = \mathbf{X}^{T}$, diagonalization is conducted:

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \mathbf{X}^{T}\mathbf{A}\mathbf{X} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{-1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 0 & -2 & 0 \\ -2 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ 1 & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

EXAMPLE: Diagonalization of Symmetrix Matrix

- To diagonalize

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}, \quad \text{with eigenvalues 4, 1, 1}$$

- Recall previously:

For $\lambda = 4$,

$$\mathbf{e}_1 = k\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\underbrace{\text{for any k} \neq 0}_{\text{Orthogonal to eigenvector of different eigenvalue}}$$

For $\lambda = 1, 1$,

$$\mathbf{e}_2 = k_1\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = k_2\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

$$\underbrace{\text{for any nonzero } k_1, k_2}_{\text{Eigenvectors of eigenevalue with multiplicity are LI but not orthogonal to each other}}$$

EXAMPLE: Diagonalization of Symmetrix Matrix

- Obtained an orthonormal eigenbasis for R^3: $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$

$$\hat{\mathbf{e}}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}, \hat{\mathbf{e}}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\\0\end{bmatrix}, \hat{\mathbf{e}}_3 = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\1\\-2\end{bmatrix}$$

$$\mathrm{X} = \begin{bmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}}\end{bmatrix}, \quad \underbrace{\mathrm{X}^{-1} = \mathrm{X}^{\mathrm{T}}}, \quad \mathrm{X}^{-1}\mathrm{A}\mathrm{X} = \mathrm{D} = \begin{bmatrix}4 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$$

Easily obtained X^{-1} as X^{T}

Quadratic Forms. Transformation to Principal Axes

- Definition: a **quadratic form** Q in the components $x_1, \dots,$ x_n of a vector $\mathbf{x}$ is a sum n^2 of terms, namely,

(7)
$$Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k$$
$$\begin{aligned} &= a_{11}x_1^{\,2} + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ &\quad + a_{21}x_2x_1 + a_{22}x_2^{\,2} + \cdots + a_{2n}x_2x_n \\ &\quad + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\quad + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^{\,2}. \end{aligned}$$

- $\mathbf{A} = [a_{jk}]$ is called the **coefficient matrix** of the form.
- May assume $\mathbf{A}$ symmetric matrix in general.
  Why? See next Example.

EXAMPLE 5
Quadratic Form. Symmetric Coefficient Matrix

Let

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

$$= 3x_1^{\ 2} + 4x_1x_2 + 6x_2x_1 + 2x_2^{\ 2}$$

$$= 3x_1^{\ 2} + 10x_1x_2 + 2x_2^{\ 2}.$$

- Here 4 + 6 = 10 = 5 + 5 so we can write in terms of a symmetric matrix.

$$\mathbf{x}^{\mathsf{T}}\mathbf{C}\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \overbrace{\begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix}}^{\mathbf{C}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 3x_1^{\ 2} + 5x_1x_2 + 5x_2x_1 + 2x_2^{\ 2}$$

$$= 3x_1^{\ 2} + 10x_1x_2 + 2x_2^{\ 2}.$$

- Another way to show $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{C}\mathbf{x}$:
  - Decompose **A** into a symmetric and skew-symmetric part:

$$\mathbf{A} = \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix}}_{\mathbf{C}} + \underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{\mathbf{S}}$$

  - where **C** is symmetric and **S** is skew-symmetric matrix.
  - Let $\mathbf{x}^T\mathbf{S}\mathbf{x}=\alpha$, then $\alpha=\alpha^T=\mathbf{x}^T\underbrace{\mathbf{S}^T}_{\mathbf{S}^T=-\mathbf{S}}\mathbf{x}= -\mathbf{x}^T\mathbf{S}\mathbf{x}=-\alpha \rightarrow \alpha=0$
  - Hence, $\mathbf{x}^T\mathbf{S}\mathbf{x}=0$ for all skew-symmetric matrix, and we have

$$\mathbf{x}^T\mathbf{A}\mathbf{x}= \mathbf{x}^T(\mathbf{C}+ \mathbf{S})\mathbf{x}= \mathbf{x}^T\mathbf{C}\mathbf{x}$$

associating with the symmetric matrix **C** only.

- Hence, assume **symmetric A** in the **quadratic form**
- By previous results, *symmetric* matrix **A** readily induces orthonormal basis of eigenvectors (for both distinct and multiple eigenvalue case).
- Using the corresponding eigenvectors as column vectors to form matrix **X** , **X** is orthogonal matrix, $\mathbf{X}^{-1} = \mathbf{X}^\mathsf{T}$.
- With (5) we have $\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$, or $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1} = \mathbf{X}\mathbf{D}\mathbf{X}^\mathsf{T}$.
- Substitute $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^\mathsf{T}$ into (7) gives

$$Q = \mathbf{x}^\mathsf{T}\mathbf{X}\mathbf{D}\mathbf{X}^\mathsf{T}\mathbf{x} \tag{8}$$

- Define $\mathbf{y} = \mathbf{X}^\mathsf{T}\mathbf{x}$, or equivalently, $\mathbf{x} = \mathbf{X}\mathbf{y}$, Q becomes simply,

$$Q = \mathbf{y}^\mathsf{T}\mathbf{D}\mathbf{y} = \lambda_1 y_1^{\,2} + \lambda_2 y_2^{\,2} + \ldots + \lambda_n y_n^{\,2}. \tag{10}$$

Theorem 5

Principal Axes Theorem

The substitution **y= X**$^{\mathsf{T}}$**x,** or **x** = **Xy,** transforms a quadratic form,

~ n^2 terms

$$Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k \qquad (a_{kj} = a_{jk})$$

to the principal axes form, or **canonical form,**

n terms

$$Q = \mathbf{y}^{\mathsf{T}}\mathbf{D}\mathbf{y} = \sum_{j=1}^{n} \lambda_j y_j^2$$

where **y= X**$^{\mathsf{T}}$**x** and $\lambda_1, \ldots, \lambda_n$ are the (distinct or multiple) eigenvalues of the (symmetric) matrix **A**. Here, **X** is an orthogonal matrix with corresponding eigenvectors $\mathbf{x}_1, \ldots,$ $\mathbf{x}_n$, respectively, as column vectors.

EXAMPLE 6

Transformation to Principal Axes. Conic Sections

Find out what type of conic section the following quadratic form represents and transform it to principal axes:

(Q) $$Q = 17x_1^{\,2} - 30x_1x_2 + 17x_2^{\,2} = 128.$$

Solution.

- We first write $Q = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$, where

$$\mathbf{A} = \begin{bmatrix} 17 & -15 \\ -15 & 17 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

- Characteristic equation of **A**:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (17 - \lambda)^2 - 15^2 = 0$$

$$\rightarrow \text{eigenvalues } \lambda_1 = 2,\ \lambda_2 = 32.$$

EXAMPLE 6 **(continued)**

- Hence, from (10), we have

$$Q = 2y_1^{\ 2} + 32y_2^{\ 2}.$$

- From Equation **(Q)**, we have $Q = 128$, or $2y_1^2 + 32y_2^2 = 128$, that is,

$$\frac{y_1^{\ 2}}{8^2} + \frac{y_2^{\ 2}}{2^2} = 1.$$

- Equation **(Q)** thus represents a ellipse with major radius of 8 along the y_1-axis and minor radius of 2 in the y_2-axis.

- Want to know the direction of the principal axes y_1 and y_2 in the x_1x_2-coordinates as well?
  -- We need to determine orthonormal eigenvectors of **A**!

EXAMPLE 6 (continued)

- Eigenvectors:

Orthogonal eigenvectors

For $\lambda_1 = 2$, $\quad \mathbf{e}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

For $\lambda_2 = 32$, $\quad \mathbf{e}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Orthonormal eigenvectors

$$\hat{\mathbf{e}}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\hat{\mathbf{e}}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

- Form $\mathbf{X}$ using orthonormal vectors:

Orthogonal matrix

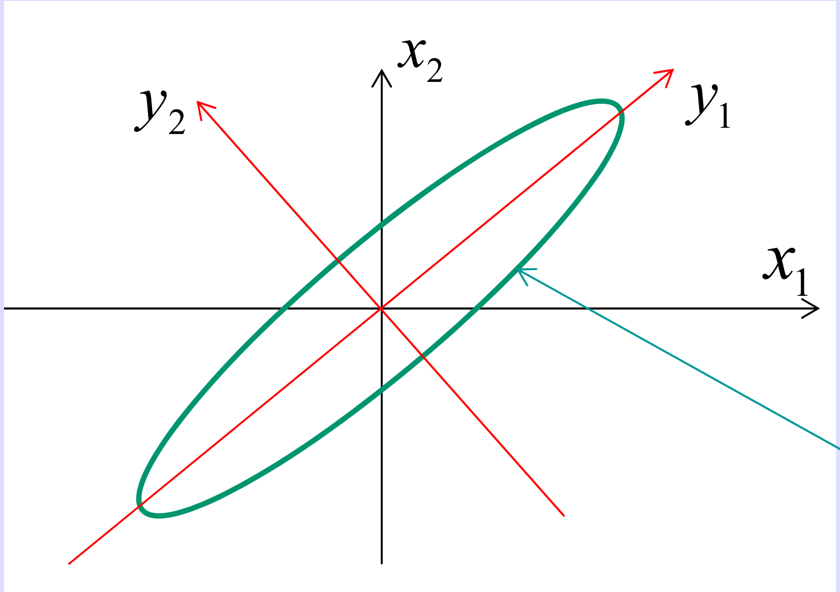
$$\mathbf{X} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \longrightarrow \mathbf{X}^{-1} = \mathbf{X}^{\mathrm{T}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

EXAMPLE 6 (continued)

Solution. Hence

$$\mathbf{x} = \mathbf{X}\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \qquad \begin{aligned} x_1 &= y_1/\sqrt{2} - y_2/\sqrt{2} \\ x_2 &= y_1/\sqrt{2} + y_2/\sqrt{2}. \end{aligned}$$

which yields a 45° rotation.



$$\left\{ \begin{aligned} Q &= 2y_1^2 + 32y_2^2 = 128 \\ Q &= 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128. \end{aligned} \right.$$

- Given the principal axes form, the **canonical form**,

$$Q = \mathbf{y}^{\mathsf{T}}\mathbf{D}\mathbf{y} = \sum_{j=1}^{n} \lambda_j y_j^2$$

  - For $\lambda_1, \ldots, \lambda_n$ all positive, the resulting conic section is ellipse, as in previous Example 6.
  - For some $\lambda_1, \ldots, \lambda_n$ negative, the resulting conic sections is hyperbola, e.g.,

$$\frac{y_1^2}{a^2} - \frac{y_2^2}{b^2} = 1$$