# vv214: Cayley-Hamilton Theorem. Symmetric matrices. Quadratic forms. Singular Value Decomposition.

Dr.Olga Danilkina

UM-SJTU Joint Institute

July 25, 2019

1. Ranking problem.
2. Cayley-Hamilton Theorem and its applications.
3. Orthogonaly diagonalizable matrices.

# Ranking Problem

* Consider the results of a tournament

$$
\begin{tikzcd}
A \arrow[r] \arrow[d] & B \arrow[d] \arrow[dl] \\
C \arrow[r] & D \arrow[ul]
\end{tikzcd}
$$

where "$A \to B$" means $A$ defeated $B$.

# Ranking Problem

* Consider the results of a tournament

$$
\begin{tikzcd}
A \arrow[r] \arrow[d] & B \arrow[d] \arrow[dl] \\
C \arrow[r] & D \arrow[ul]
\end{tikzcd}
$$

where "$A \to B$" means $A$ defeated $B$.

How to rank the players fairly?

# Ranking Problem

* Consider the results of a tournament

$$\begin{tikzcd} A \arrow[r] \arrow[d] & B \arrow[d] \arrow[dl] \\ C \arrow[r] & D \arrow[ul] \end{tikzcd}$$

where "$A \to B$" means $A$ defeated $B$.

How to rank the players fairly?

* The Page-Brin idea:

# Ranking Problem

* Consider the results of a tournament

$$\begin{tikzcd} A \arrow[r] \arrow[d] & B \arrow[d] \arrow[dl] \\ C \arrow[r] & D \arrow[ul] \end{tikzcd}$$

where "$A \to B$" means $A$ defeated $B$.

How to rank the players fairly?

* The Page-Brin idea:it should *be worth more* to defeat a *better player*

# Ranking Problem

* Consider the results of a tournament

$$\begin{tikzcd}
A \arrow[r] \arrow[d] & B \arrow[d] \arrow[dl] \\
C \arrow[r] & D \arrow[ul]
\end{tikzcd}$$

where "$A \rightarrow B$" means $A$ defeated $B$.

How to rank the players fairly?

* The Page-Brin idea:it should *be worth more* to defeat a *better player* $\Leftarrow$ based on Perron (1907)-Frobenius (1912) Theorem

# Ranking Problem

* Consider the results of a tournament

$$\begin{tikzcd} A \arrow[r] \arrow[d] & B \arrow[d] \arrow[dl] \\ C \arrow[r] & D \arrow[ul] \end{tikzcd}$$

where "$A \to B$" means $A$ defeated $B$.

How to rank the players fairly?

* The Page-Brin idea:it should *be worth more* to defeat a *better player* $\Leftarrow$ based on Perron (1907)-Frobenius (1912) Theorem

How do we know who is better before ranking them?

# Ranking Problem

* Consider the results of a tournament

$$
\begin{tikzcd}
A \arrow[r] \arrow[d] & B \arrow[d] \arrow[dl] \\
C \arrow[r] & D \arrow[ul]
\end{tikzcd}
$$

where "$A \to B$" means $A$ defeated $B$.

How to rank the players fairly?

* The Page-Brin idea:it should *be worth more* to defeat a *better player* $\Leftarrow$ based on Perron (1907)-Frobenius (1912) Theorem

How do we know who is better before ranking them?

* Define recursion!

# Ranking Problem

* Give everyone the initial score of 1

$$\bar{x}_0 = \begin{array}{cc} \left( \begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array} \right) & \begin{array}{c} A \\ B \\ C \\ D \end{array} \end{array}$$

# Ranking Problem

* Give everyone the initial score of 1

$$\bar{x}_0 = \begin{array}{cc} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} & \begin{array}{c} A \\ B \\ C \\ D \end{array} \end{array}$$

* Define for all $n \geq 0$

$$\bar{x}_{n+1} = A\bar{x}_n,$$

where

$$A = \begin{array}{cc} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

# Ranking Problem

$$\bar{x}_1 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{array}{c} \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} \end{array} \begin{array}{c} A \\ B \\ C \\ D \end{array}$$

# Ranking Problem

$$\bar{x}_1 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} \begin{array}{c} A \\ B \\ C \\ D \end{array}$$

$$\bar{x}_2 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix} \begin{array}{c} A \\ B \\ C \\ D \end{array}$$

# Ranking Problem

$$\bar{x}_1 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

$$\bar{x}_2 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

$$\bar{x}_3 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 2 \\ 3 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

# Ranking Problem

$$\bar{x}_1 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

$$\bar{x}_2 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

$$\bar{x}_3 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 2 \\ 3 \end{pmatrix} \begin{matrix} A \\ B \\ C \\ D \end{matrix}$$

*The (n+1)th score of a player A is the sum of the nth scores of the players that the player A defeated.*

# Ranking Problem

$$\bar{x}_5 = \begin{pmatrix} 8 \\ 6 \\ 3 \\ 5 \end{pmatrix}, \bar{x}_{10} = \begin{pmatrix} 35 \\ 34 \\ 21 \\ 26 \end{pmatrix}, \bar{x}_{100} = \begin{pmatrix} 1037 \\ 933 \\ 547 \\ 731 \end{pmatrix}$$

# Ranking Problem

$$\bar{x}_5 = \begin{pmatrix} 8 \\ 6 \\ 3 \\ 5 \end{pmatrix}, \bar{x}_{10} = \begin{pmatrix} 35 \\ 34 \\ 21 \\ 26 \end{pmatrix}, \bar{x}_{100} = \begin{pmatrix} 1037 \\ 933 \\ 547 \\ 731 \end{pmatrix}$$

Is $A > B > D > C$? $\Rightarrow$

# Ranking Problem

$$\bar{x}_5 = \begin{pmatrix} 8 \\ 6 \\ 3 \\ 5 \end{pmatrix}, \bar{x}_{10} = \begin{pmatrix} 35 \\ 34 \\ 21 \\ 26 \end{pmatrix}, \bar{x}_{100} = \begin{pmatrix} 1037 \\ 933 \\ 547 \\ 731 \end{pmatrix}$$

Is $A > B > D > C$? $\Rightarrow$ analyze $\bar{x}_n = A^n \bar{x}_0$ as $n \to \infty$

# Ranking Problem

$$\bar{x}_5 = \begin{pmatrix} 8 \\ 6 \\ 3 \\ 5 \end{pmatrix}, \bar{x}_{10} = \begin{pmatrix} 35 \\ 34 \\ 21 \\ 26 \end{pmatrix}, \bar{x}_{100} = \begin{pmatrix} 1037 \\ 933 \\ 547 \\ 731 \end{pmatrix}$$

Is $A > B > D > C$? $\Rightarrow$ analyze $\bar{x}_n = A^n \bar{x}_0$ as $n \to \infty$

**Theorem (Perron-Frobenius):** There exists a *largest positive* eigenvalue $\lambda_{PF}$ for a nonnegative matrix $A$ such that the rescaled system

$$\bar{x}_n = \left(\frac{1}{\lambda_{PF}} A\right)^n \bar{x}_0$$

converges to an equilibrium state $\bar{x}_\infty$.

$$\bar{x}_\infty = \bar{x}_{\infty+1} = \frac{1}{\lambda_{PF}} A \bar{x}_\infty \Rightarrow A\bar{x}_\infty = \lambda_{PF} \bar{x}_\infty$$

*The equilibrium state is the eigenvector associated with $\lambda_{PF}$!!!*

# Ranking Problem

The largest positive eigenvalue is

$$\lambda_{PF} = 1.3953369...$$

and

$$\bar{x}_\infty = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{pmatrix} 0.321... \\ 0.288... \\ 0.165... \\ 0.230... \end{pmatrix} \begin{array}{c} A \\ B \\ C \\ D \end{array}$$

# Ranking Problem

The largest positive eigenvalue is

$$\lambda_{PF} = 1.3953369...$$

and

$$\bar{x}_\infty = \begin{array}{c} \\ \\ \\ \\ \end{array}\left(\begin{array}{c} 0.321... \\ 0.288... \\ 0.165... \\ 0.230... \end{array}\right)\begin{array}{c} A \\ B \\ C \\ D \end{array}$$

$$A > B > D > C$$

# Cayley-Hamilton Theorem: Intro

1. Let $g(t) = a_0 + a_1 t + \ldots + a_k t^k$ and $D = diag(d_1, \ldots, d_n)$

# Cayley-Hamilton Theorem: Intro

1. Let $g(t) = a_0 + a_1 t + \ldots + a_k t^k$ and $D = diag(d_1, \ldots, d_n)$

$$\Rightarrow g(D) = diag(g(d_1), \ldots, g(d_n))$$

# Cayley-Hamilton Theorem: Intro

1. Let $g(t) = a_0 + a_1 t + \ldots + a_k t^k$ and $D = diag(d_1, \ldots, d_n)$

$$\Rightarrow g(D) = diag(g(d_1), \ldots, g(d_n))$$

Let $f_D(\lambda)$ be a char poly $\Rightarrow$

# Cayley-Hamilton Theorem: Intro

1. Let $g(t) = a_0 + a_1 t + \ldots + a_k t^k$ and $D = diag(d_1, \ldots, d_n)$

$$\Rightarrow g(D) = diag(g(d_1), \ldots, g(d_n))$$

Let $f_D(\lambda)$ be a char poly $\Rightarrow f_D(D) = \begin{pmatrix} f_D(d_1) & & \\ & \ddots & \\ & & f_D(d_n) \end{pmatrix} =$

# Cayley-Hamilton Theorem: Intro

1. Let $g(t) = a_0 + a_1 t + \ldots + a_k t^k$ and $D = diag(d_1, \ldots, d_n)$

$$\Rightarrow g(D) = diag(g(d_1), \ldots, g(d_n))$$

Let $f_D(\lambda)$ be a char poly $\Rightarrow f_D(D) = \begin{pmatrix} f_D(d_1) & & \\ & \ddots & \\ & & f_D(d_n) \end{pmatrix} =$

$$= \begin{pmatrix} \prod_{i=1}^n (d_1 - d_i) & & \\ & \ddots & \\ & & \prod_{i=1}^n (d_n - d_i) \end{pmatrix} =$$

# Cayley-Hamilton Theorem: Intro

1. Let $g(t) = a_0 + a_1 t + \ldots + a_k t^k$ and $D = diag(d_1, \ldots, d_n)$

$$\Rightarrow g(D) = diag(g(d_1), \ldots, g(d_n))$$

Let $f_D(\lambda)$ be a char poly $\Rightarrow f_D(D) = \begin{pmatrix} f_D(d_1) & & \\ & \ddots & \\ & & f_D(d_n) \end{pmatrix} =$

$$= \begin{pmatrix} \prod_{i=1}^{n}(d_1 - d_i) & & \\ & \ddots & \\ & & \prod_{i=1}^{n}(d_n - d_i) \end{pmatrix} = \underbrace{0}_{\text{zero matrix}}$$

# Cayley-Hamilton Theorem: Intro

1. Let $g(t) = a_0 + a_1 t + \ldots + a_k t^k$ and $D = diag(d_1, \ldots, d_n)$

$$\Rightarrow g(D) = diag(g(d_1), \ldots, g(d_n))$$

Let $f_D(\lambda)$ be a char poly $\Rightarrow f_D(D) = \begin{pmatrix} f_D(d_1) & & \\ & \ddots & \\ & & f_D(d_n) \end{pmatrix} =$

$$= \begin{pmatrix} \prod_{i=1}^n (d_1 - d_i) & & \\ & \ddots & \\ & & \prod_{i=1}^n (d_n - d_i) \end{pmatrix} = \underbrace{0}_{\text{zero matrix}}$$

$$\boxed{f_D(D) = 0}$$

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow$

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow D = S^{-1}AS$ ($A$ is diagonalizable)

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow D = S^{-1}AS$ ($A$ is diagonalizable)

$$g(D) = a_0 I_n + a_1 D + \ldots + a_k D^k =$$

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow D = S^{-1}AS$ ($A$ is diagonalizable)

$$g(D) = a_0 I_n + a_1 D + \ldots + a_k D^k =$$

$$= a_0 S^{-1} S + a_1 S^{-1} AS + \ldots + a_k S^{-1} A^k S =$$

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow D = S^{-1}AS$ ($A$ is diagonalizable)

$$g(D) = a_0 I_n + a_1 D + \ldots + a_k D^k =$$

$$= a_0 S^{-1} S + a_1 S^{-1} A S + \ldots + a_k S^{-1} A^k S = S^{-1}(a_0 I_n + a_1 A + \ldots + A^k) S$$

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow D = S^{-1}AS$ ($A$ is diagonalizable)

$$g(D) = a_0 I_n + a_1 D + \ldots + a_k D^k =$$

$$= a_0 S^{-1} S + a_1 S^{-1} A S + \ldots + a_k S^{-1} A^k S = S^{-1}(a_0 I_n + a_1 A + \ldots + A^k) S$$

$$\Rightarrow g(D) = S^{-1} g(A) S \Rightarrow$$

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow D = S^{-1}AS$ ($A$ is diagonalizable)

$$g(D) = a_0 I_n + a_1 D + \ldots + a_k D^k =$$

$$= a_0 S^{-1} S + a_1 S^{-1} A S + \ldots + a_k S^{-1} A^k S = S^{-1}(a_0 I_n + a_1 A + \ldots + A^k) S$$

$$\Rightarrow g(D) = S^{-1} g(A) S \Rightarrow g(D) \sim g(A)$$

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow D = S^{-1}AS$ ($A$ is diagonalizable)

$$g(D) = a_0 I_n + a_1 D + \ldots + a_k D^k =$$

$$= a_0 S^{-1}S + a_1 S^{-1}AS + \ldots + a_k S^{-1}A^k S = S^{-1}(a_0 I_n + a_1 A + \ldots + A^k)S$$

$$\Rightarrow g(D) = S^{-1}g(A)S \Rightarrow g(D) \sim g(A)$$

$$A \sim D \Rightarrow f_A(\lambda) = f_D(\lambda) \Rightarrow f_A(D) = f_D(D) = 0$$

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow D = S^{-1}AS$ ($A$ is diagonalizable)

$$g(D) = a_0 I_n + a_1 D + \ldots + a_k D^k =$$

$$= a_0 S^{-1}S + a_1 S^{-1}AS + \ldots + a_k S^{-1}A^k S = S^{-1}(a_0 I_n + a_1 A + \ldots + A^k)S$$

$$\Rightarrow g(D) = S^{-1}g(A)S \Rightarrow g(D) \sim g(A)$$

$$A \sim D \Rightarrow f_A(\lambda) = f_D(\lambda) \Rightarrow f_A(D) = f_D(D) = 0$$

$$f_A(A) = S\,\underbrace{f_A(D)}_{0}\,S^{-1} = 0$$

# Cayley-Hamilton Theorem: Intro

2. Let $A$ be similar to $D \Rightarrow D = S^{-1}AS$ ($A$ is diagonalizable)

$$g(D) = a_0 I_n + a_1 D + \ldots + a_k D^k =$$

$$= a_0 S^{-1}S + a_1 S^{-1}AS + \ldots + a_k S^{-1}A^k S = S^{-1}(a_0 I_n + a_1 A + \ldots + A^k)S$$

$$\Rightarrow g(D) = S^{-1}g(A)S \Rightarrow g(D) \sim g(A)$$

$$A \sim D \Rightarrow f_A(\lambda) = f_D(\lambda) \Rightarrow f_A(D) = f_D(D) = 0$$

$$f_A(A) = S\,\underbrace{f_A(D)}_{0}\,S^{-1} = 0$$

$$\boxed{f_A(A) = 0}$$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D =$

## Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} =$$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} =$$

## Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow$$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

2. Find a sequence of diagonalizable matrices that converges to $A$.

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

2. Find a sequence of diagonalizable matrices that converges to $A$.

$$\{B_m\},$$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

2. Find a sequence of diagonalizable matrices that converges to $A$.

$$\{B_m\},\ B_m = \begin{pmatrix} 1 & 1 \\ 0 & 1 + \frac{1}{m} \end{pmatrix}$$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

2. Find a sequence of diagonalizable matrices that converges to $A$.

$$\{B_m\},\ B_m = \begin{pmatrix} 1 & 1 \\ 0 & 1 + \frac{1}{m} \end{pmatrix}$$

$B_m$ has distinct eigenvalues and hence, diagonalizable

$$\Rightarrow$$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

2. Find a sequence of diagonalizable matrices that converges to $A$.

$$\{B_m\},\ B_m = \begin{pmatrix} 1 & 1 \\ 0 & 1 + \frac{1}{m} \end{pmatrix}$$

$B_m$ has distinct eigenvalues and hence, diagonalizable

$$\Rightarrow f_{B_m}(B_m) = 0$$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

2. Find a sequence of diagonalizable matrices that converges to $A$.

$$\{B_m\},\ B_m = \begin{pmatrix} 1 & 1 \\ 0 & 1 + \frac{1}{m} \end{pmatrix}$$

$B_m$ has distinct eigenvalues and hence, diagonalizable

$$\Rightarrow f_{B_m}(B_m) = 0$$

$B_m \to A \Rightarrow$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

2. Find a sequence of diagonalizable matrices that converges to $A$.

$$\{B_m\},\ B_m = \begin{pmatrix} 1 & 1 \\ 0 & 1 + \frac{1}{m} \end{pmatrix}$$

$B_m$ has distinct eigenvalues and hence, diagonalizable

$$\Rightarrow f_{B_m}(B_m) = 0$$

$$B_m \to A \Rightarrow f_{B_m} \to f_A \text{ and determinant is a continuous function}$$

# Cayley-Hamilton Theorem: Intro

1. Is $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalizable?

If $D = S^{-1}AS$ then $f_A(\lambda) = (\lambda - 1)^2 = f_D(\lambda) \Rightarrow D = I_2$

$$A = SDS^{-1} = SIS^{-1} = I \Rightarrow \text{contradiction}$$

2. Find a sequence of diagonalizable matrices that converges to $A$.

$$\{B_m\},\ B_m = \begin{pmatrix} 1 & 1 \\ 0 & 1 + \frac{1}{m} \end{pmatrix}$$

$B_m$ has distinct eigenvalues and hence, diagonalizable

$$\Rightarrow f_{B_m}(B_m) = 0$$

$$B_m \to A \Rightarrow f_{B_m} \to f_A \text{ and determinant is a continuous function}$$

$$f_A(A) = \lim_{m\to\infty} f_{B_m}(B_m) = 0$$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A) I_n = 0$$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\,A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n \times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A) I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

## Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A) I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow$$

## Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A =$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A = (4A - 5I)A =$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A = (4A - 5I)A = 4A^2 - 5A =$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A) I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A = (4A - 5I)A = 4A^2 - 5A = 4(4A - 5I) - 5A =$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\,A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A = (4A-5I)A = 4A^2-5A = 4(4A-5I)-5A = 11A-20I$ etc.

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A) I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A = (4A - 5I)A = 4A^2 - 5A = 4(4A - 5I) - 5A = 11A - 20I$ etc.

2. If $A$ is orthogonal with $\det A = 1$ and $\lambda_1 = -\frac{1}{2} - \frac{\sqrt{3}}{2} i$, then

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A) I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A = (4A-5I)A = 4A^2 - 5A = 4(4A-5I) - 5A = 11A - 20I$ etc.

2. If $A$ is orthogonal with $\det A = 1$ and $\lambda_1 = -\frac{1}{2} - \frac{\sqrt{3}}{2} i$, then $\lambda_2 = -\frac{1}{2} + \frac{\sqrt{3}}{2} i$ and

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A) I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A = (4A-5I)A = 4A^2 - 5A = 4(4A-5I) - 5A = 11A - 20I$ etc.

2. If $A$ is orthogonal with $\det A = 1$ and $\lambda_1 = -\frac{1}{2} - \frac{\sqrt{3}}{2} i$, then $\lambda_2 = -\frac{1}{2} + \frac{\sqrt{3}}{2} i$ and $\lambda_3 = 1$.

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

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\,A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2A = (4A-5I)A = 4A^2-5A = 4(4A-5I)-5A = 11A-20I$ etc.

2. If $A$ is orthogonal with $\det A = 1$ and $\lambda_1 = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$, then $\lambda_2 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\lambda_3 = 1$.
$f_A(\lambda) = (-\lambda)^3 + (-1+1)(-\lambda)^2 + \det A \Rightarrow -A^3 + I = 0 \Rightarrow A^3 = I$

## Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A = (4A-5I)A = 4A^2-5A = 4(4A-5I)-5A = 11A-20I$ etc.

2. If $A$ is orthogonal with $\det A = 1$ and $\lambda_1 = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$, then $\lambda_2 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\lambda_3 = 1$.

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

$$A^{214} =$$

## Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A) I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2 A = (4A - 5I)A = 4A^2 - 5A = 4(4A - 5I) - 5A = 11A - 20I \quad$ etc.

2. If $A$ is orthogonal with $\det A = 1$ and $\lambda_1 = -\frac{1}{2} - \frac{\sqrt{3}}{2} i$, then $\lambda_2 = -\frac{1}{2} + \frac{\sqrt{3}}{2} i$ and $\lambda_3 = 1$.

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

$$A^{214} = (A^3)^{71} A =$$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\,A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2A = (4A-5I)A = 4A^2-5A = 4(4A-5I)-5A = 11A-20I$ etc.

2. If $A$ is orthogonal with $\det A = 1$ and $\lambda_1 = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$, then $\lambda_2 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\lambda_3 = 1$.

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

$$A^{214} = (A^3)^{71}A = I^{71}A =$$

# Cayley-Hamilton Theorem

**Theorem:** Any $A \in M_{n\times n}(\mathbb{K})$ satisfies its own characteristic equation, i.e.

$$f_A(A) = (-A)^n + (tr\, A)(-A)^{n-1} + \ldots + (\det A)I_n = 0$$

**Examples:**

1. $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \Rightarrow f_A(\lambda) = \lambda^2 - 4\lambda + 5 = 0 \Rightarrow \lambda_{1,2} = 2 \pm i$

$$f_A(A) = A^2 - 4A + 5I = 0 \Rightarrow A^2 = 4A - 5I$$

$A^3 = A^2A = (4A-5I)A = 4A^2-5A = 4(4A-5I)-5A = 11A-20I$ etc.

2. If $A$ is orthogonal with $\det A = 1$ and $\lambda_1 = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$, then $\lambda_2 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\lambda_3 = 1$.

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

$$A^{214} = (A^3)^{71}A = I^{71}A = A$$

# Cayley-Hamilton Theorem: Order Reduction

3. Order Reduction

Represent $g(t) = h(t)f_A(t) + r(t) \Rightarrow$

# Cayley-Hamilton Theorem: Order Reduction

3. Order Reduction

Represent $g(t) = h(t)f_A(t) + r(t) \Rightarrow g(\lambda) = r(\lambda)$

# Cayley-Hamilton Theorem: Order Reduction

3. Order Reduction

Represent $g(t) = h(t)f_A(t) + r(t) \Rightarrow g(\lambda) = r(\lambda)$

$$g(A) = h(A)f_A(A) + r(A) \Rightarrow$$

# Cayley-Hamilton Theorem: Order Reduction

3. Order Reduction

Represent $g(t) = h(t)f_A(t) + r(t) \Rightarrow g(\lambda) = r(\lambda)$

$$g(A) = h(A)f_A(A) + r(A) \Rightarrow g(A) = r(A)$$

# Cayley-Hamilton Theorem: Order Reduction

3. Order Reduction

Represent $g(t) = h(t)f_A(t) + r(t) \Rightarrow g(\lambda) = r(\lambda)$

$$g(A) = h(A)f_A(A) + r(A) \Rightarrow g(A) = r(A)$$

Let $g(A) = A^5 + 2A^4 - A^3 + A^2 - 2A + I$ with $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$

# Cayley-Hamilton Theorem: Order Reduction

3. Order Reduction

Represent $g(t) = h(t)f_A(t) + r(t) \Rightarrow g(\lambda) = r(\lambda)$

$$g(A) = h(A)f_A(A) + r(A) \Rightarrow g(A) = r(A)$$

Let $g(A) = A^5 + 2A^4 - A^3 + A^2 - 2A + I$ with $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$

$$f_A(\lambda) = \lambda^2 + 5\lambda + 6$$

## Cayley-Hamilton Theorem: Order Reduction

3. Order Reduction

Represent $g(t) = h(t)f_A(t) + r(t) \Rightarrow g(\lambda) = r(\lambda)$

$$g(A) = h(A)f_A(A) + r(A) \Rightarrow g(A) = r(A)$$

Let $g(A) = A^5 + 2A^4 - A^3 + A^2 - 2A + I$ with $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$

$$f_A(\lambda) = \lambda^2 + 5\lambda + 6$$

$$\Rightarrow g(\lambda) = (\lambda^3 - 3\lambda^2 + 8\lambda - 21)(\lambda^2 + 5\lambda + 6) + \underbrace{65\lambda + 127}_{r(\lambda)}$$

# Cayley-Hamilton Theorem: Order Reduction

3. Order Reduction

Represent $g(t) = h(t)f_A(t) + r(t) \Rightarrow g(\lambda) = r(\lambda)$

$$g(A) = h(A)f_A(A) + r(A) \Rightarrow g(A) = r(A)$$

Let $g(A) = A^5 + 2A^4 - A^3 + A^2 - 2A + I$ with $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$

$$f_A(\lambda) = \lambda^2 + 5\lambda + 6$$

$$\Rightarrow g(\lambda) = (\lambda^3 - 3\lambda^2 + 8\lambda - 21)(\lambda^2 + 5\lambda + 6) + \underbrace{65\lambda + 127}_{r(\lambda)}$$

$$g(A) = 65A + 127I$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Let a complex-valued function $g(t)$ be analytic in some region of the complex plane $\Rightarrow$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Let a complex-valued function $g(t)$ be analytic in some region of the complex plane $\Rightarrow g(t) = \sum_{k=0}^{\infty} a_k t^k$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Let a complex-valued function $g(t)$ be analytic in some region of the complex plane $\Rightarrow g(t) = \sum_{k=0}^{\infty} a_k t^k$

$$g(t) = h(t) f_A(t) + r(t) \Rightarrow$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Let a complex-valued function $g(t)$ be analytic in some region of the complex plane $\Rightarrow g(t) = \sum_{k=0}^{\infty} a_k t^k$

$$g(t) = h(t) f_A(t) + r(t) \Rightarrow g(\lambda_i) = r(\lambda_i) =$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Let a complex-valued function $g(t)$ be analytic in some region of the complex plane $\Rightarrow g(t) = \sum_{k=0}^{\infty} a_k t^k$

$$g(t) = h(t) f_A(t) + r(t) \Rightarrow g(\lambda_i) = r(\lambda_i) = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Let a complex-valued function $g(t)$ be analytic in some region of the complex plane $\Rightarrow g(t) = \sum_{k=0}^{\infty} a_k t^k$

$$g(t) = h(t) f_A(t) + r(t) \Rightarrow g(\lambda_i) = r(\lambda_i) = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

$$g(A) = \sum_{k=0}^{n-1} b_k A^k$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$,

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

Let $g(t) = \sin t \Rightarrow$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

Let $g(t) = \sin t \Rightarrow \begin{array}{l} \sin(\lambda_1) = b_0 + b_1\lambda_1 \\ \sin(\lambda_2) = b_0 + b_1\lambda_2 \end{array}$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

Let $g(t) = \sin t \Rightarrow \begin{array}{l} \sin(\lambda_1) = b_0 + b_1\lambda_1 \\ \sin(\lambda_2) = b_0 + b_1\lambda_2 \end{array}$

$$\Rightarrow \begin{array}{rl} b_0 = & 3\sin(-2) - 2\sin(-3) \\ b_1 = & \sin(-2) - \sin(-3) \end{array}$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

Let $g(t) = \sin t \Rightarrow \begin{array}{l} \sin(\lambda_1) = b_0 + b_1\lambda_1 \\ \sin(\lambda_2) = b_0 + b_1\lambda_2 \end{array}$

$$\Rightarrow \begin{array}{rc} b_0 = & 3\sin(-2) - 2\sin(-3) \\ b_1 = & \sin(-2) - \sin(-3) \end{array}$$

$$\sin A = g(A) = r(A) = b_0 I + b_1 A$$

$$\sin A = (3\sin(-2) - 2\sin(-3))I + (\sin(-2) - \sin(-3))A$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

Let $g(t) = \sin t \Rightarrow \begin{array}{l} \sin(\lambda_1) = b_0 + b_1\lambda_1 \\ \sin(\lambda_2) = b_0 + b_1\lambda_2 \end{array}$

$$\Rightarrow \begin{array}{rcc} b_0 & = & 3\sin(-2) - 2\sin(-3) \\ b_1 & = & \sin(-2) - \sin(-3) \end{array}$$

$$\sin A = g(A) = r(A) = b_0 I + b_1 A$$

$$\sin A = (3\sin(-2) - 2\sin(-3))I + (\sin(-2) - \sin(-3))A$$

$$\sin \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix} =$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

Let $g(t) = \sin t \Rightarrow \begin{array}{l} \sin(\lambda_1) = b_0 + b_1\lambda_1 \\ \sin(\lambda_2) = b_0 + b_1\lambda_2 \end{array}$

$$\Rightarrow \begin{array}{rcc} b_0 & = & 3\sin(-2) - 2\sin(-3) \\ b_1 & = & \sin(-2) - \sin(-3) \end{array}$$

$$\sin A = g(A) = r(A) = b_0 I + b_1 A$$

$$\sin A = (3\sin(-2) - 2\sin(-3))I + (\sin(-2) - \sin(-3))A$$

$$\sin \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} \sin(-3) & \sin(-2) - \sin(-3) \\ 0 & \sin(-2) \end{pmatrix}$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k,$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

$$e^{-2x} = b_0 - 2b_1,\ e^{-3x} = b_0 - 3b_1$$

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

$$e^{-2x} = b_0 - 2b_1, \; e^{-3x} = b_0 - 3b_1$$

$$\begin{aligned} b_0 &= 3e^{-2x} - 2e^{-3x} \\ b_1 &= e^{-2x} - e^{-3x} \end{aligned}$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

$$e^{-2x} = b_0 - 2b_1, \; e^{-3x} = b_0 - 3b_1$$

$$\begin{aligned} b_0 &= \; 3e^{-2x} - 2e^{-3x} \\ b_1 &= \; e^{-2x} - e^{-3x} \end{aligned}$$

$$e^{Ax} =$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

$$e^{-2x} = b_0 - 2b_1,\ e^{-3x} = b_0 - 3b_1$$

$$\begin{aligned} b_0 &= \ 3e^{-2x} - 2e^{-3x} \\ b_1 &= \ \ e^{-2x} - e^{-3x} \end{aligned}$$

$$e^{Ax} = (3e^{-2x} - 2e^{-3x})I +$$

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

$$e^{-2x} = b_0 - 2b_1,\ e^{-3x} = b_0 - 3b_1$$

$$\begin{aligned} b_0 &= \ 3e^{-2x} - 2e^{-3x} \\ b_1 &= \ e^{-2x} - e^{-3x} \end{aligned}$$

$$e^{Ax} = (3e^{-2x} - 2e^{-3x})I + (e^{-2x} - e^{-3x})A$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

$$e^{-2x} = b_0 - 2b_1, \; e^{-3x} = b_0 - 3b_1$$

$$\begin{aligned} b_0 &= \; 3e^{-2x} - 2e^{-3x} \\ b_1 &= \; e^{-2x} - e^{-3x} \end{aligned}$$

$$e^{Ax} = (3e^{-2x} - 2e^{-3x})I + (e^{-2x} - e^{-3x})A$$

$$exp\begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix} x =$$

# Cayley-Hamilton Theorem: Analytic Functions of a Matrix

4. Matrix Exponential

$$g(t) = e^{tx} \Rightarrow e^{Ax} = \sum_{k=0}^{n-1} b_k A^k, \quad e^{\lambda_i x} = \sum_{k=0}^{n-1} b_k \lambda_i^k$$

Let $A = \begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix}$, $f_A(\lambda) = \lambda^2 + 5\lambda + 6 \Rightarrow r(t) = b_0 + b_1 t$

$$e^{-2x} = b_0 - 2b_1, \; e^{-3x} = b_0 - 3b_1$$

$$\begin{aligned} b_0 &= \; 3e^{-2x} - 2e^{-3x} \\ b_1 &= \; e^{-2x} - e^{-3x} \end{aligned}$$

$$e^{Ax} = (3e^{-2x} - 2e^{-3x})I + (e^{-2x} - e^{-3x})A$$

$$exp\begin{pmatrix} -3 & 1 \\ 0 & -2 \end{pmatrix} x = \begin{pmatrix} e^{-3x} & e^{-2x} - e^{-2x} \\ 0 & e^{-2x} \end{pmatrix}$$

# Cayley-Hamilton Theorem: Minimal Polynomial

**Definition:** The smallest degree polynomial $m_A(t) \neq 0$ such that $m_A(A) = 0$ is called the minimal polynomial of $A$.

# Cayley-Hamilton Theorem: Minimal Polynomial

**Definition:** The smallest degree polynomial $m_A(t) \neq 0$ such that $m_A(A) = 0$ is called the minimal polynomial of $A$.

The minimal polynomial of

$$D = \begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 2 \end{pmatrix}$$

is $m_D(t) = (t-1)(t-2)$.

# Cayley-Hamilton Theorem: Minimal Polynomial

**Definition:** The smallest degree polynomial $m_A(t) \neq 0$ such that $m_A(A) = 0$ is called the minimal polynomial of $A$.

The minimal polynomial of

$$D = \begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 2 \end{pmatrix}$$

is $m_D(t) = (t-1)(t-2)$.

Is the polynomial $t - 1$ minimal for $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$?

## Cayley-Hamilton Theorem: Minimal Polynomial

**Definition:** The smallest degree polynomial $m_A(t) \neq 0$ such that $m_A(A) = 0$ is called the minimal polynomial of $A$.

The minimal polynomial of

$$D = \begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 2 \end{pmatrix}$$

is $m_D(t) = (t-1)(t-2)$.

Is the polynomial $t - 1$ minimal for $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$?no!

# Cayley-Hamilton Theorem: Minimal Polynomial

**Definition:** The smallest degree polynomial $m_A(t) \neq 0$ such that $m_A(A) = 0$ is called the minimal polynomial of $A$.
The minimal polynomial of

$$D = \begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 2 \end{pmatrix}$$

is $m_D(t) = (t-1)(t-2)$.

Is the polynomial $t-1$ minimal for $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$?no!

Cayley-Hamilton Theorem $\Rightarrow \forall A \in M_{n\times n}(\mathbb{K}) \quad m_A | f_A$

# Cayley-Hamilton Theorem: Minimal Polynomial

**Definition:** The smallest degree polynomial $m_A(t) \neq 0$ such that $m_A(A) = 0$ is called the minimal polynomial of $A$.
The minimal polynomial of

$$D = \begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 2 \end{pmatrix}$$

is $m_D(t) = (t-1)(t-2)$.
Is the polynomial $t-1$ minimal for $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$?no!
Cayley-Hamilton Theorem $\Rightarrow \forall A \in M_{n\times n}(\mathbb{K}) \quad m_A | f_A$

$$\Rightarrow m_A(t) = (t-1)^2$$

# Cayley-Hamilton Theorem: Minimal Polynomial

**Definition:** The smallest degree polynomial $m_A(t) \neq 0$ such that $m_A(A) = 0$ is called the minimal polynomial of $A$.

The minimal polynomial of

$$D = \begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 2 \end{pmatrix}$$

is $m_D(t) = (t-1)(t-2)$.

Is the polynomial $t-1$ minimal for $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$?no!

Cayley-Hamilton Theorem $\Rightarrow \forall A \in M_{n\times n}(\mathbb{K}) \quad m_A | f_A$

$$\Rightarrow m_A(t) = (t-1)^2$$

**Remark:** Let $g(t)$ be a polynomial with coefficients from $\mathbb{K}$. $g$ has multiple roots iff $\gcd(g, g')$ is not a constant.

# Orthogonally Diagonalizable Matrices

1. Question: What are the conditions on a square matrix to grantee that it is diagonalizable?

# Orthogonally Diagonalizable Matrices

1. Question: What are the conditions on a square matrix to grantee that it is diagonalizable?
   Answer: The existence of an eigenbasis.

# Orthogonally Diagonalizable Matrices

1. Question: What are the conditions on a square matrix to grantee that it is diagonalizable?
   Answer: The existence of an eigenbasis.
2. Question: For which matrices is there an orthonormal eigenbasis?

# Orthogonally Diagonalizable Matrices

1. Question: What are the conditions on a square matrix to grantee that it is diagonalizable?
   Answer: The existence of an eigenbasis.
2. Question: For which matrices is there an orthonormal eigenbasis?
   Answer: For which matrices is there an orthogonal matrix $S$ such that $S^{-1}AS = S^{T}AS$ is diagonal?

# Orthogonally Diagonalizable Matrices

1. Question: What are the conditions on a square matrix to grantee that it is diagonalizable?
   Answer: The existence of an eigenbasis.
2. Question: For which matrices is there an orthonormal eigenbasis?
   Answer: For which matrices is there an orthogonal matrix $S$ such that $S^{-1}AS = S^{T}AS$ is diagonal?

**Definition:** A matrix $A$ is orthogonally diagonalizable if there exists an orthogonal matrix $S$ such that

$$S^{-1}AS = S^{T}AS$$

is diagonal.

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \qquad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \quad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \quad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

$$A \quad \text{is diagonalizable}$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix},$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \quad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

$$A \quad \text{is diagonalizable}$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \qquad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

$$A \quad \text{is diagonalizable}$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow \bar{u}_1 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix},$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \quad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

$$A \quad \text{is diagonalizable}$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow \bar{u}_1 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \quad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

$$A \quad \text{is diagonalizable}$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow \bar{u}_1 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

$$\Rightarrow \bar{u}_1,\ \bar{u}_2 \text{ is the orthonormal basis}$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

$$A \quad \text{is diagonalizable}$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow \bar{u}_1 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

$$\Rightarrow \bar{u}_1,\ \bar{u}_2 \text{ is the orthonormal basis}$$

$$S = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \quad \text{is orthogonal and}$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \quad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

$$A \quad \text{is diagonalizable}$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow \bar{u}_1 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

$$\Rightarrow \bar{u}_1,\ \bar{u}_2 \text{ is the orthonormal basis}$$

$$S = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \quad \text{is orthogonal and} \quad D = S^{-1}AS = \begin{pmatrix} 3 & 0 \\ 0 & 8 \end{pmatrix}$$

# Example 1

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 7 \end{pmatrix} \quad A \quad \text{is symmetric}$$

$$|A - \lambda I| = 0 \Rightarrow (4-\lambda)(7-\lambda) - 4 = 0 \Rightarrow \quad \underbrace{\lambda_1 = 3,\ \lambda_2 = 8}_{\text{2 different eigenvalues}}$$

$$A \quad \text{is diagonalizable}$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow \bar{u}_1 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

$$\Rightarrow \bar{u}_1,\ \bar{u}_2 \text{ is the orthonormal basis}$$

$$S = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \quad \text{is orthogonal and} \quad D = S^{-1}AS = \begin{pmatrix} 3 & 0 \\ 0 & 8 \end{pmatrix}$$

$A$ is orthogonally diagonalizable

# Example 2

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

# Example 2

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad |A - \lambda I| = 0 \Rightarrow \lambda^2(3 - \lambda) = 0 \Rightarrow \lambda_{1,2} = 0,\ \lambda_3 = 3$$

# Example 2

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad |A - \lambda I| = 0 \Rightarrow \lambda^2(3 - \lambda) = 0 \Rightarrow \lambda_{1,2} = 0,\ \lambda_3 = 3$$

$$\lambda_{1,2} = 0\colon\ \bar{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},\ \lambda_3 = 3\colon\ \bar{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

# Example 2

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad |A - \lambda I| = 0 \Rightarrow \lambda^2(3-\lambda) = 0 \Rightarrow \lambda_{1,2} = 0,\ \lambda_3 = 3$$

$$\lambda_{1,2} = 0:\ \bar{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},\ \lambda_3 = 3:\ \bar{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\Rightarrow \bar{u}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ 2 \end{pmatrix}, \bar{u}_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

# Example 2

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad |A - \lambda I| = 0 \Rightarrow \lambda^2(3-\lambda) = 0 \Rightarrow \lambda_{1,2} = 0,\ \lambda_3 = 3$$

$$\lambda_{1,2} = 0:\ \bar{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},\ \lambda_3 = 3:\ \bar{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\Rightarrow \bar{u}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ 2 \end{pmatrix}, \bar{u}_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

$$S = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 2 & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix},$$

# Example 2

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad |A - \lambda I| = 0 \Rightarrow \lambda^2(3-\lambda) = 0 \Rightarrow \lambda_{1,2} = 0,\ \lambda_3 = 3$$

$$\lambda_{1,2} = 0:\ \bar{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},\ \lambda_3 = 3:\ \bar{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\Rightarrow \bar{u}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ 2 \end{pmatrix}, \bar{u}_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

$$S = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 2 & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix},\ D = S^{-1}AS = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

# Example 2

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad |A - \lambda I| = 0 \Rightarrow \lambda^2(3-\lambda) = 0 \Rightarrow \lambda_{1,2} = 0,\ \lambda_3 = 3$$

$$\lambda_{1,2} = 0:\ \bar{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \bar{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},\ \lambda_3 = 3:\ \bar{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\Rightarrow \bar{u}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \bar{u}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ 2 \end{pmatrix}, \bar{u}_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

$$S = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 2 & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix},\ D = S^{-1}AS = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$A$ is orthogonally diagonalizable

# Useful Statements

**Lemma 1:** If $A$ is orthogonally diagonalizable, then $A^T = A$.

# Useful Statements

**Lemma 1:** If $A$ is orthogonally diagonalizable, then $A^T = A$.

**Lemma 2:** Let $A^T = A$. If $\bar{v}_1$, $\bar{v}_2$ are eigenvectors of $A$ with distinct eigenvalues $\lambda_1, \lambda_2$, then $\bar{v}_1 \perp \bar{v}_2$.

# Useful Statements

**Lemma 1:** If $A$ is orthogonally diagonalizable, then $A^T = A$.

**Lemma 2:** Let $A^T = A$. If $\bar{v}_1$, $\bar{v}_2$ are eigenvectors of $A$ with distinct eigenvalues $\lambda_1$, $\lambda_2$, then $\bar{v}_1 \perp \bar{v}_2$.

**Lemma 3:** A *symmetric* matrix $A_{n\times n}$ has $n$ *real* eigenvalues counted with their algebraic multiplicities.

# Useful Statements

**Lemma 1:** If $A$ is orthogonally diagonalizable, then $A^T = A$.

**Lemma 2:** Let $A^T = A$. If $\bar{v}_1$, $\bar{v}_2$ are eigenvectors of $A$ with distinct eigenvalues $\lambda_1$, $\lambda_2$, then $\bar{v}_1 \perp \bar{v}_2$.

**Lemma 3:** A *symmetric* matrix $A_{n\times n}$ has $n$ *real* eigenvalues counted with their algebraic multiplicities.

**Theorem (Spectral Theorem):** A matrix $A$ is orthogonally diagonalizable iff $A$ is symmetric ($A^T = A$).

# Adjoint Operators

**Definition:** Let $T \colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star \colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V\, \forall w \in W$$

# Adjoint Operators

**Definition:** Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V \, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

# Adjoint Operators

**Definition:** Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V \, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$

# Adjoint Operators

**Definition:** Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V\, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.

# Adjoint Operators

**Definition:** Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V \, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.
3. By the Riesz Representation Theorem, there exists a unique element $b \in V\colon (Tv, w) = (v, b)$

# Adjoint Operators

**Definition:** Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V \, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.
3. By the Riesz Representation Theorem, there exists a unique element $b \in V\colon (Tv, w) = (v, b)$
4. Denote $b = T^\star w$. Since $b$ exists, so $T^\star w$ exists as well.

# Adjoint Operators

**Definition:** Let $T: V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star : W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V \, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.
3. By the Riesz Representation Theorem, there exists a unique element $b \in V$: $(Tv, w) = (v, b)$
4. Denote $b = T^\star w$. Since $b$ exists, so $T^\star w$ exists as well.
5. Therefore, there exists a map $w \to T^\star w$.

# Adjoint Operators

**Definition:** Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V\, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.
3. By the Riesz Representation Theorem, there exists a unique element $b \in V\colon (Tv, w) = (v, b)$
4. Denote $b = T^\star w$. Since $b$ exists, so $T^\star w$ exists as well.
5. Therefore, there exists a map $w \to T^\star w$.

**Example:** Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1)$$

# Adjoint Operators

**Definition:** Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V\, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.
3. By the Riesz Representation Theorem, there exists a unique element $b \in V\colon (Tv, w) = (v, b)$
4. Denote $b = T^\star w$. Since $b$ exists, so $T^\star w$ exists as well.
5. Therefore, there exists a map $w \to T^\star w$.

**Example:** Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1)$$

$$(T\bar{x}, \bar{y}) = ((x_2 + 3x_3, 2x_1), (y_1, y_2)) =$$

# Adjoint Operators

**Definition:** Let $T: V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star: W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V \, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.
3. By the Riesz Representation Theorem, there exists a unique element $b \in V$: $(Tv, w) = (v, b)$
4. Denote $b = T^\star w$. Since $b$ exists, so $T^\star w$ exists as well.
5. Therefore, there exists a map $w \to T^\star w$.

**Example:** Let $T: \mathbb{R}^3 \to \mathbb{R}^2$

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1)$$

$$(T\bar{x}, \bar{y}) = ((x_2 + 3x_3, 2x_1), (y_1, y_2)) = x_2 y_1 + 3x_3 y_1 + 2x_1 y_2$$

# Adjoint Operators

**Definition:** Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V\,\forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.
3. By the Riesz Representation Theorem, there exists a unique element $b \in V\colon (Tv, w) = (v, b)$
4. Denote $b = T^\star w$. Since $b$ exists, so $T^\star w$ exists as well.
5. Therefore, there exists a map $w \to T^\star w$.

**Example:** Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1)$$

$$(T\bar{x}, \bar{y}) = ((x_2 + 3x_3, 2x_1), (y_1, y_2)) = x_2y_1 + 3x_3y_1 + 2x_1y_2$$

$$= ((x_1, x_2, x_3), (2y_2, y_1, 3y_1))$$

# Adjoint Operators

**Definition:** Let $T\colon V \to W$ be linear, i.e. $T \in L(V, W)$.
The adjoint of $T$ is the operator $T^\star\colon W \to V$ such that

$$(Tv, w) = (v, T^\star w) \quad \forall v \in V \, \forall w \in W$$

**Remark:** The adjoint $T^\star$ is well defined:

1. Fix $w \in W$. Define a map $V \to \mathbb{K}$ such that $v \to (Tv, w)$
2. The map $v \to (Tv, w)$ is linear $\Rightarrow$ it is a linear functional.
3. By the Riesz Representation Theorem, there exists a unique element $b \in V\colon (Tv, w) = (v, b)$
4. Denote $b = T^\star w$. Since $b$ exists, so $T^\star w$ exists as well.
5. Therefore, there exists a map $w \to T^\star w$.

**Example:** Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1)$$

$$(T\bar{x}, \bar{y}) = ((x_2 + 3x_3, 2x_1), (y_1, y_2)) = x_2 y_1 + 3x_3 y_1 + 2x_1 y_2$$

$$= ((x_1, x_2, x_3), (2y_2, y_1, 3y_1)) \Rightarrow T^\star \bar{y} = (2y_2, y_1, 3y_1)$$

# Properties of Adjoint Operators

1. $T \in L(V, W) \Rightarrow T^\star \in L(W, V)$

$$(v, T^\star(w_1 + w_2)) = (Tv, w_1 + w_2) = (Tv, w_1) + (Tv, w_2)$$

$$= (v, T^\star w_1) + (v, T^\star w_2) = (v, T^\star w_1 + T^\star w_2)$$

$$(v, T^\star(\lambda w)) = (Tv, \lambda w) = \bar{\lambda}(Tv, w) = \bar{\lambda}(v, T^\star w) = (v, \lambda T^\star w)$$

2. $(T_1 + T_2)^\star = T_1^\star + T_2^\star \quad \forall S, T \in L(V, W)$
3. $(\lambda T)^\star = \bar{\lambda} T^\star \quad \forall \lambda \in \mathbb{K} \, \forall T \in L(V, W)$
4. $(T^\star)^\star = T \quad \forall T \in L(V, W)$

$$(w, (T^\star)^\star v) = (T^\star w, v) = \overline{(v, T^\star w)} = \overline{(Tv, w)} = (w, Tv) \, \forall v \in V$$

5. $I^\star = I$
6. $(T_1 T_2)^\star = T_2^\star T_1^\star \quad \forall T_1 \in L(W, U),\ T_2 \in L(V, W)$

# Properties of Adjoint Operators

7. $Ker\ T^\star = (Im\ T)^\perp$

$$w \in Ker\ T^\star \Leftrightarrow T^\star w = 0 \Leftrightarrow (v, T^\star w) = 0 \quad \forall v \in V$$

$$\Leftrightarrow (Tv, w) = 0 \quad \forall v \in V \Leftrightarrow w \in (Im\ T)^\perp$$

8. The matrix of the adjoint $T^\star$ w.r.t orthonormal bases $e_1, \ldots, e_m \in V$; $f_1, \ldots f_n \in W$ is the conjugate transpose of the matrix of $T$.

$$A_T = (Te_1\ Te_2 \ldots Te_m) \quad Te_k \in W,\ f_1, \ldots, f_n \text{ is orthonormal}$$

$$\Rightarrow Te_k = (Te_k, f_1)f_1 + \ldots + (Te_k, f_n)f_n \Rightarrow (A_T)_{jk} = (Te_k, f_j)$$

$$A_{T^\star} = (T^\star f_1\ T^\star f_2 \ldots T^\star f_m) \quad T^\star f_k \in V,\ e_1, \ldots, e_m \text{ is orthonormal}$$

$$\Rightarrow T^\star f_k = (T^\star f_k, e_1)e_1 + \ldots + (T^\star f_k, e_m)e_m$$

$$\Rightarrow (A_{T^\star})_{jk} = (T^\star f_k, e_j) = \overline{(e_j, T^\star f_k)} = \overline{(Te_j, f_k)}$$

## Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v,\, w \in V$$

## Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v,\, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v,\, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T\colon \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v,\, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T: \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $\lambda ||v||^2 =$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T: \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $\lambda ||v||^2 = (\lambda v, v) =$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v,\ w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T: \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $\lambda||v||^2 = (\lambda v, v) = (Tv, v) =$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $\lambda||v||^2 = (\lambda v, v) = (Tv, v) = (v, Tv) =$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v,\, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T: \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $\lambda||v||^2 = (\lambda v, v) = (Tv, v) = (v, Tv) = (v, \lambda v) =$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v,\, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $\lambda||v||^2 = (\lambda v, v) = (Tv, v) = (v, Tv) = (v, \lambda v) = \bar{\lambda}(v, v) = \bar{\lambda}||v||^2$

3. $T$ is self-adjoint iff $(Tv, v) \in \mathbb{R} \quad \forall v \in V$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V,V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $$\lambda||v||^2 = (\lambda v, v) = (Tv, v) = (v, Tv) = (v, \lambda v) = \bar{\lambda}(v, v) = \bar{\lambda}||v||^2$$

3. $T$ is self-adjoint iff $(Tv, v) \in \mathbb{R} \quad \forall v \in V$

   $$(Tv.v) - \overline{(Tv, v)} =$$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $\lambda||v||^2 = (\lambda v, v) = (Tv, v) = (v, Tv) = (v, \lambda v) = \bar{\lambda}(v, v) = \bar{\lambda}||v||^2$
3. $T$ is self-adjoint iff $(Tv, v) \in \mathbb{R} \quad \forall v \in V$

   $(Tv.v) - \overline{(Tv, v)} = (Tv, v) - (v, Tv) =$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v, w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $$\lambda||v||^2 = (\lambda v, v) = (Tv, v) = (v, Tv) = (v, \lambda v) = \bar{\lambda}(v, v) = \bar{\lambda}||v||^2$$

3. $T$ is self-adjoint iff $(Tv, v) \in \mathbb{R} \quad \forall v \in V$

   $$(Tv.v) - \overline{(Tv, v)} = (Tv, v) - (v, Tv) = (Tv, v) - (T^\star v, v)$$

   $$=$$

# Self-Adjoint Operators

**Definition:** The operator $T \in L(V, V)$ is called self-adjoint if $T^\star = T$:

$$(Tv, w) = (v, Tw) \quad \forall v,\ w \in V$$

Hermitian=self-adjoint

**Remarks:**

1. Let $T : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ be defined by the matrix $\begin{pmatrix} 1 & a \\ 2 & 3 \end{pmatrix} \Rightarrow T$ is self-adjoint iff $a = 2$, that is, its matrix is symmetric.
2. Every eigenvalue of a self-adjoint operator is real.

   $$\lambda||v||^2 = (\lambda v, v) = (Tv, v) = (v, Tv) = (v, \lambda v) = \bar{\lambda}(v, v) = \bar{\lambda}||v||^2$$

3. $T$ is self-adjoint iff $(Tv, v) \in \mathbb{R} \quad \forall v \in V$

   $$(Tv.v) - \overline{(Tv, v)} = (Tv, v) - (v, Tv) = (Tv, v) - (T^\star v, v)$$

   $$= ((T - T^\star)v, v)$$

# Normal Operators

**Definition:** An operator on an inner product space is called normal if it commutes with its adjoint, i.e $TT^{\star} = T^{\star}T$.

# Normal Operators

**Definition:** An operator on an inner product space is called normal if it commutes with its adjoint, i.e $TT^\star = T^\star T$.

**Remarks:**

1. Every self-adjoint operator is normal.
2. $T$ is normal iff $||Tv|| = ||T^\star v|| \, \forall v$

   $T$ is normal $\Leftrightarrow$

# Normal Operators

**Definition:** An operator on an inner product space is called normal if it commutes with its adjoint, i.e $TT^\star = T^\star T$.

**Remarks:**

1. Every self-adjoint operator is normal.
2. $T$ is normal iff $||Tv|| = ||T^\star v|| \, \forall v$

$$T \text{ is normal} \Leftrightarrow T^\star T - TT^\star = 0 \Leftrightarrow$$

# Normal Operators

**Definition:** An operator on an inner product space is called normal if it commutes with its adjoint, i.e $TT^\star = T^\star T$.

**Remarks:**

1. Every self-adjoint operator is normal.
2. $T$ is normal iff $||Tv|| = ||T^\star v|| \, \forall v$

$$T \text{ is normal} \Leftrightarrow T^\star T - TT^\star = 0 \Leftrightarrow ((T^\star T - TT^\star)v, v) = 0$$
$$\Leftrightarrow$$

# Normal Operators

**Definition:** An operator on an inner product space is called normal if it commutes with its adjoint, i.e $TT^\star = T^\star T$.

**Remarks:**

1. Every self-adjoint operator is normal.
2. $T$ is normal iff $||Tv|| = ||T^\star v|| \, \forall v$

$$T \text{ is normal} \Leftrightarrow T^\star T - TT^\star = 0 \Leftrightarrow ((T^\star T - TT^\star)v, v) = 0$$
$$\Leftrightarrow (T^\star Tv, v) = (TT^\star v, v) \Leftrightarrow$$

# Normal Operators

**Definition:** An operator on an inner product space is called normal if it commutes with its adjoint, i.e $TT^\star = T^\star T$.

**Remarks:**

1. Every self-adjoint operator is normal.
2. $T$ is normal iff $||Tv|| = ||T^\star v|| \, \forall v$

$$T \text{ is normal} \Leftrightarrow T^\star T - TT^\star = 0 \Leftrightarrow ((T^\star T - TT^\star)v, v) = 0$$
$$\Leftrightarrow (T^\star Tv, v) = (TT^\star v, v) \Leftrightarrow ||Tv||^2 = ||T^\star v||^2$$

3. Let $Tv = \lambda v$. $T$ is normal $\Leftrightarrow$ $T - \lambda I$ is also normal.
4. Let $Tv = \lambda v$. Then $T^\star v = \bar{\lambda} v$

$$0 = ||(T - \lambda I)v|| = ||(T - \lambda I)^\star v|| = ||(T^\star - \bar{\lambda} I)v|| \Rightarrow T^\star v = \bar{\lambda} v$$

5. Suppose $T \in L(V, V)$ is normal. Then eigenvectors of $T$ corresponding to distinct eigenvalues are orthogonal.

$$Tu = \alpha u, \; Tv = \beta v \Rightarrow T^\star v = \bar{\beta} v$$

$$(\alpha - \beta)(u, v) = \alpha(u, v) - \beta(u, v) = (Tu, v) - (u, T^\star v) = 0$$

# Complex Spectral Theorem

Let $\mathbb{K} = \mathbb{C}$ and $T \in L(V, V)$. Then the following are equivalent:
(a) $T$ is normal.
(b) $V$ has an orthonormal basis consisting of eigenvectors of $T$.
(c) $T$ has a diagonal matrix with respect to some orthonormal basis of $V$.

## Quadratic Forms: Motivation

Consider a function $q(\bar{x}) = q(x_1, x_2) \colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2$$

# Quadratic Forms: Motivation

Consider a function $q(\bar{x}) = q(x_1, x_2) \colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2$$

$q(0,0) = 0 \Rightarrow$

# Quadratic Forms: Motivation

Consider a function $q(\bar{x}) = q(x_1, x_2)\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2$$

$q(0,0) = 0 \Rightarrow$ Is $q(0,0)$ a global maximum or a global minimum or neither?

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2 =$$

# Quadratic Forms: Motivation

Consider a function $q(\bar{x}) = q(x_1, x_2) \colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2$$

$q(0,0) = 0 \Rightarrow$ Is $q(0,0)$ a global maximum or a global minimum or neither?

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2 = 8x_1^2 - 2x_1x_2 - 2x_1x_2 + 5x_2^2 =$$

# Quadratic Forms: Motivation

Consider a function $q(\bar{x}) = q(x_1,\, x_2)\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2$$

$q(0,0) = 0 \Rightarrow$ Is $q(0,0)$ a global maximum or a global minimum or neither?

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2 = 8x_1^2 - 2x_1x_2 - 2x_1x_2 + 5x_2^2 =$$

$$= x_1(8x_1 - 2x_2) + x_2(-2x_1 + 5x_2) =$$

# Quadratic Forms: Motivation

Consider a function $q(\bar{x}) = q(x_1, x_2) \colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2$$

$q(0,0) = 0 \Rightarrow$ Is $q(0,0)$ a global maximum or a global minimum or neither?

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2 = 8x_1^2 - 2x_1x_2 - 2x_1x_2 + 5x_2^2 =$$

$$= x_1(8x_1 - 2x_2) + x_2(-2x_1 + 5x_2) =$$

$$= \left( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 8x_1 - 2x_2 \\ -2x_1 + 5x_2 \end{pmatrix} \right) =$$

# Quadratic Forms: Motivation

Consider a function $q(\bar{x}) = q(x_1, x_2)\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2$$

$q(0,0) = 0 \Rightarrow$ Is $q(0,0)$ a global maximum or a global minimum or neither?

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2 = 8x_1^2 - 2x_1x_2 - 2x_1x_2 + 5x_2^2 =$$

$$= x_1(8x_1 - 2x_2) + x_2(-2x_1 + 5x_2) =$$

$$= \left( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 8x_1 - 2x_2 \\ -2x_1 + 5x_2 \end{pmatrix} \right) = (\bar{x}, A\bar{x}) =$$

# Quadratic Forms: Motivation

Consider a function $q(\bar{x}) = q(x_1,\, x_2)\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2$$

$q(0,0) = 0 \Rightarrow$ Is $q(0,0)$ a global maximum or a global minimum or neither?

$$q(\bar{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2 = 8x_1^2 - 2x_1x_2 - 2x_1x_2 + 5x_2^2 =$$

$$= x_1(8x_1 - 2x_2) + x_2(-2x_1 + 5x_2) =$$

$$= \left( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} 8x_1 - 2x_2 \\ -2x_1 + 5x_2 \end{pmatrix} \right) = (\bar{x},\, A\bar{x}) = \bar{x}^T A\bar{x}$$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1, \ \bar{v}_2 \text{ for } A$$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A$$

$\lambda_1 = 9,\ \lambda_2 = 4 \Rightarrow$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A$$

$$\lambda_1 = 9,\ \lambda_2 = 4 \Rightarrow \bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix},$$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A$$

$$\lambda_1 = 9,\ \lambda_2 = 4 \Rightarrow \bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A$$

$$\lambda_1 = 9,\ \lambda_2 = 4 \Rightarrow \bar{v}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$\bar{x} = \alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2 \Rightarrow$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A$$

$$\lambda_1 = 9,\ \lambda_2 = 4 \Rightarrow \bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\bar{x} = \alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2 \Rightarrow q(\bar{x}) = (\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2,\ A(\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2)) =$$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A$$

$$\lambda_1 = 9,\ \lambda_2 = 4 \Rightarrow \bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\bar{x} = \alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2 \Rightarrow q(\bar{x}) = (\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2,\ A(\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2)) =$$

$$= (\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2,\ 9\alpha_1 \bar{v}_1 + 4\alpha_2 \bar{v}_2) =$$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A$$

$$\lambda_1 = 9,\ \lambda_2 = 4 \Rightarrow \bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\bar{x} = \alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2 \Rightarrow q(\bar{x}) = (\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2,\ A(\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2)) =$$

$$= (\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2,\ 9\alpha_1 \bar{v}_1 + 4\alpha_2 \bar{v}_2) = 9\alpha_1^2 + 4\alpha_2^2$$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A$$

$$\lambda_1 = 9,\ \lambda_2 = 4 \Rightarrow \bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\bar{x} = \alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2 \Rightarrow q(\bar{x}) = (\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2,\ A(\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2)) =$$

$$= (\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2,\ 9\alpha_1 \bar{v}_1 + 4\alpha_2 \bar{v}_2) = 9\alpha_1^2 + 4\alpha_2^2 \geq 0$$

# Quadratic Forms: Motivation

The matrix

$$A = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix}$$

is symmetric *by the construction* $\Rightarrow$Spectral Theorem

$$\exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A$$

$$\lambda_1 = 9,\ \lambda_2 = 4 \Rightarrow \bar{v}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \bar{v}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\bar{x} = \alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2 \Rightarrow q(\bar{x}) = (\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2,\ A(\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2)) =$$

$$= (\alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2,\ 9\alpha_1 \bar{v}_1 + 4\alpha_2 \bar{v}_2) = 9\alpha_1^2 + 4\alpha_2^2 \geq 0$$

$$q(0,0) = 0 \text{ is the global minimum}$$

# Quadratic Forms: Definition

**Definition:** A function $q(x_1, \ldots, x_n)\colon \mathbb{R}^n \to \mathbb{R}$ is called a quadratic form if it is a linear combination of products $x_i x_j$.

# Quadratic Forms: Definition

**Definition:** A function $q(x_1, \ldots, x_n)\colon \mathbb{R}^n \to \mathbb{R}$ is called a quadratic form if it is a linear combination of products $x_i x_j$.

**Remark 1:** A quadratic form can be represented as

$$q(\bar{x}) = (\bar{x},\, A\bar{x}) = \bar{x}^T A\bar{x},$$

*$A$ is a symmetric matrix, the representation is unique.*

# Quadratic Forms: Definition

**Definition:** A function $q(x_1, \ldots, x_n) \colon \mathbb{R}^n \to \mathbb{R}$ is called a quadratic form if it is a linear combination of products $x_i x_j$.
**Remark 1:** A quadratic form can be represented as

$$q(\bar{x}) = (\bar{x},\, A\bar{x}) = \bar{x}^T A\bar{x},$$

*A is a symmetric matrix, the representation is unique.*
**Remark 2:** Let

$$Q_n = \{\text{quadratic}\, q(x_1, \ldots, x_n) \colon \mathbb{R}^n \to \mathbb{R}\}$$

$$\forall q \in Q_n \quad q = c_1 x_1^2 + c_2 x_1 x_2 + c_3 x_2^2 + c_4 x_3^2 + c_5 x_1 x_3 + c_6 x_2 x_3 + \ldots$$

$$\Rightarrow \{x_i x_j\},\ i, j = 1 \ldots n, \text{ is the basis for } Q_n$$

$$n + (n-1) + (n-2) + \ldots + 2 + 1 = \frac{n(n+1)}{2} \quad \text{basis elements}$$

# Theorem

If

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

is a quadratic form with a *symmetric matrix* $A$ and

$$\mathfrak{B} = \{\bar{v}_1, \ldots, \bar{v}_n\}$$

is the eigenbasis for $A$ with the associated eigenvalues $\lambda_1, \ldots, \lambda_n$, then

$$q(\bar{x}) = \lambda_1 \alpha_1^2 + \ldots + \lambda_n \alpha_n^2,$$

where $\alpha_1, \ldots, \alpha_n$ are the coordinates of $\bar{x}$ in $\mathfrak{B}$.

# Positive Definite Quadratic Forms

**Definition:** A quadratic form

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

is said to be positive definite if

$$q(\bar{x}) > 0 \quad \forall \bar{x} \neq \bar{0},\ \bar{x} \in \mathbb{R}^n$$

If $q(\bar{x}) \geq 0 \, \forall \bar{x} \in \mathbb{R}^n$ the $q$ is positive semi-definite.
**Example:** Let $q(\bar{x}) = ||A\bar{x}||^2$, $A_{n \times m}$.

# Positive Definite Quadratic Forms

**Definition:** A quadratic form

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

is said to be positive definite if

$$q(\bar{x}) > 0 \quad \forall \bar{x} \neq \bar{0},\ \bar{x} \in \mathbb{R}^n$$

If $q(\bar{x}) \geq 0\, \forall \bar{x} \in \mathbb{R}^n$ the $q$ is positive semi-definite.
**Example:** Let $q(\bar{x}) = ||A\bar{x}||^2$, $A_{n\times m}$.

$$q(\bar{x}) = (A\bar{x},\, A\bar{x}) = (A\bar{x})^T A\bar{x} = \bar{x}^T \underbrace{A^T A}_{\text{symmetric}} \bar{x}$$

$$q(\bar{x}) = ||A\bar{x}||^2 \Rightarrow q(\bar{x}) \geq 0 \quad \forall \bar{x} \in \mathbb{R}^n \Rightarrow q \text{ is positive semidefinite}$$

$$q(\bar{x}) = 0 \text{ iff } \bar{x} \in Ker\ A$$

# Remarks

**Remark1:** A quadratic form

$$q(\bar{x}) = \lambda_1 \alpha_1^2 + \ldots \lambda_n \alpha_n^2$$

is positive semi-definite iff all $\lambda_1, \ldots, \lambda_n$ are positive or zero.

# Remarks

**Remark1:** A quadratic form

$$q(\bar{x}) = \lambda_1 \alpha_1^2 + \ldots \lambda_n \alpha_n^2$$

is positive semi-definite iff all $\lambda_1, \ldots, \lambda_n$ are positive or zero.

**Remark 2:** If $q(\bar{x}) = \bar{x}^T A \bar{x}$ is positive definite then

$$\det A = \lambda_1 \cdots \lambda_n > 0$$

BUT the converse is not true: Let $A_{3\times 3}$

$$\lambda_1 > 0,\ \lambda_2,\ \lambda_3 < 0 \Rightarrow \det A = \lambda_1 \lambda_2 \lambda_3 > 0$$

$$q(\bar{x}) = \underbrace{\lambda_1 \alpha_1^2}_{>0} + \underbrace{\lambda_2 \alpha_2^2}_{<0} + \underbrace{\lambda_3 \alpha_3^2}_{<0} \Rightarrow q \text{ is indefinite}$$

**Definition:** A symmetric matrix is called positive definite provided all of its eigenvalues are positive.

# Theorem

Let $A^{(m)}$ be a *principal submatrix* of a symmetric matrix $A$ obtained by omitting all rows and columns of $A$ past the $m$th. Then $A$ is positive definite iff

$$\det A^{(m)} > 0 \quad \forall m = 1, \ldots, n$$

**Example:**

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 2 & 4 \\ -2 & 4 & 3 \end{pmatrix} \quad |A^{(1)}| = 1,\ |A^{(2)}| = 2,\ |A^{(3)}| = -18$$

$A$ is not positive definite

$$B = \begin{pmatrix} 9 & -1 & 2 \\ -1 & 7 & -3 \\ 2 & -3 & 3 \end{pmatrix} \quad |B^{(1)}| = 9,\ |B^{(2)}| = 62,\ |B^{(3)}| = 89$$

$B$ is positive definite

# Motivation

Consider the equation

$$q(x_1, x_2) = 1$$

for the quadratic form

$$q\colon \mathbb{R}^2 \to \mathbb{R},\ q(x_1, x_2) = ax_1^2 + bx_2^2.$$

# Motivation

Consider the equation

$$q(x_1, x_2) = 1$$

for the quadratic form

$$q\colon \mathbb{R}^2 \to \mathbb{R},\ q(x_1, x_2) = ax_1^2 + bx_2^2.$$

1. $b > a > 0 \Rightarrow$

$$\frac{x_1^2}{(1/\sqrt{a})^2} + \frac{x_2^2}{(1/\sqrt{b})^2} = 1$$

is an ellipse with semimajor and semiminor axes $\dfrac{1}{\sqrt{a}}, \dfrac{1}{\sqrt{b}}$

# Motivation

Consider the equation

$$q(x_1, x_2) = 1$$

for the quadratic form

$$q\colon \mathbb{R}^2 \to \mathbb{R},\ q(x_1, x_2) = ax_1^2 + bx_2^2.$$

1. $b > a > 0 \Rightarrow$

$$\frac{x_1^2}{(1/\sqrt{a})^2} + \frac{x_2^2}{(1/\sqrt{b})^2} = 1$$

is an ellipse with semimajor and semiminor axes $\dfrac{1}{\sqrt{a}}, \dfrac{1}{\sqrt{b}}$

2. $a > 0,\ b < 0 \Rightarrow$

$$\frac{x_1^2}{(1/\sqrt{a})^2} - \frac{x_2^2}{(1/\sqrt{-b})^2} = 1$$

is a hyperbola.

# Theorem

Let $\lambda_1$, $\lambda_2$ be distinct eigenvalues of the matrix

$$\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

of the quadratic form

$$q(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2.$$

If $\lambda_1 \cdot \lambda_2 > 0$, then the curve

$$C \subset \mathbb{R}^2 : \; q(x_1, x_2) = 1$$

is an *ellipse*.
If $\lambda_1 \cdot \lambda_2 < 0$, the $C$ is *hyperbola*.

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^T A)^T =$$

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^T A)^T = A^T (A^T)^T =$$

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^T A)^T = A^T (A^T)^T = A^T A \Rightarrow$$

# SVD: Motivation

Let $L \colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^T A)^T = A^T (A^T)^T = A^T A \Rightarrow A^T A \text{ is symmetric}$$

$$\text{Spectral Theorem} \Rightarrow \exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A^T A$$

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^TA)^T = A^T(A^T)^T = A^TA \Rightarrow A^TA \text{ is symmetric}$$

$$\text{Spectral Theorem} \Rightarrow \exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A^TA$$

$(A\bar{v}_1, A\bar{v}_2) =$

# SVD: Motivation

Let $L \colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^T A)^T = A^T (A^T)^T = A^T A \Rightarrow A^T A \text{ is symmetric}$$

Spectral Theorem $\Rightarrow \exists$ an orthonormal basis $\bar{v}_1, \; \bar{v}_2$ for $A^T A$

$$(A\bar{v}_1, A\bar{v}_2) = (A\bar{v}_1)^T A\bar{v}_2 =$$

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^TA)^T = A^T(A^T)^T = A^TA \Rightarrow A^TA \text{ is symmetric}$$

$$\text{Spectral Theorem} \Rightarrow \exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A^TA$$

$$(A\bar{v}_1, A\bar{v}_2) = (A\bar{v}_1)^T A\bar{v}_2 = \bar{v}_1^T A^T A\bar{v}_2 =$$

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^T A)^T = A^T (A^T)^T = A^T A \Rightarrow A^T A \text{ is symmetric}$$

$$\text{Spectral Theorem} \Rightarrow \exists \text{ an orthonormal basis } \bar{v}_1,\ \bar{v}_2 \text{ for } A^T A$$

$$(A\bar{v}_1, A\bar{v}_2) = (A\bar{v}_1)^T A\bar{v}_2 = \bar{v}_1^T A^T A\bar{v}_2 = \bar{v}_1^T \lambda_2 \bar{v}_2 =$$

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^TA)^T = A^T(A^T)^T = A^TA \Rightarrow A^TA \text{ is symmetric}$$

Spectral Theorem $\Rightarrow \exists$ an orthonormal basis $\bar{v}_1,\ \bar{v}_2$ for $A^TA$

$$(A\bar{v}_1, A\bar{v}_2) = (A\bar{v}_1)^T A\bar{v}_2 = \bar{v}_1^T A^T A\bar{v}_2 = \bar{v}_1^T \lambda_2 \bar{v}_2 = \lambda_2(\bar{v}_1,\ \bar{v}_2) =$$

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^T A)^T = A^T (A^T)^T = A^T A \Rightarrow A^T A \text{ is symmetric}$$

Spectral Theorem $\Rightarrow \exists$ an orthonormal basis $\bar{v}_1,\ \bar{v}_2$ for $A^T A$

$$(A\bar{v}_1, A\bar{v}_2) = (A\bar{v}_1)^T A\bar{v}_2 = \bar{v}_1^T A^T A\bar{v}_2 = \bar{v}_1^T \lambda_2 \bar{v}_2 = \lambda_2(\bar{v}_1,\ \bar{v}_2) = 0$$

# SVD: Motivation

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, $L\bar{x} = A\bar{x}$.

$$(A^T A)^T = A^T (A^T)^T = A^T A \Rightarrow A^T A \text{ is symmetric}$$

Spectral Theorem $\Rightarrow \exists$ an orthonormal basis $\bar{v}_1,\ \bar{v}_2$ for $A^T A$

$$(A\bar{v}_1, A\bar{v}_2) = (A\bar{v}_1)^T A\bar{v}_2 = \bar{v}_1^T A^T A\bar{v}_2 = \bar{v}_1^T \lambda_2 \bar{v}_2 = \lambda_2(\bar{v}_1,\ \bar{v}_2) = 0$$

$$\Rightarrow A\bar{v}_1 \perp A\bar{v}_2$$

# SVD: Example

Let $L \colon \mathbb{R}^2 \to \mathbb{R}^2$, $L\bar{x} = A\bar{x}$, $A = \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix}$

# SVD: Example

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L\bar{x} = A\bar{x}$, $A = \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix}$

$$A^T A = \begin{pmatrix} 6 & -7 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix} = \begin{pmatrix} 85 & -30 \\ -30 & 40 \end{pmatrix}$$

# SVD: Example

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L\bar{x} = A\bar{x}$, $A = \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix}$

$$A^T A = \begin{pmatrix} 6 & -7 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix} = \begin{pmatrix} 85 & -30 \\ -30 & 40 \end{pmatrix}$$

Eigenvalues: $(85 - \lambda)(40 - \lambda) - 900 = 0 \Rightarrow \lambda_1 = 100,\ \lambda_2 = 25$

# SVD: Example

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L\bar{x} = A\bar{x}$, $A = \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix}$

$$A^T A = \begin{pmatrix} 6 & -7 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix} = \begin{pmatrix} 85 & -30 \\ -30 & 40 \end{pmatrix}$$

Eigenvalues: $(85 - \lambda)(40 - \lambda) - 900 = 0 \Rightarrow \lambda_1 = 100,\ \lambda_2 = 25$

Orthonormal eigenbasis: $\bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix} \perp \bar{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

# SVD: Example

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L\bar{x} = A\bar{x}$, $A = \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix}$

$$A^T A = \begin{pmatrix} 6 & -7 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix} = \begin{pmatrix} 85 & -30 \\ -30 & 40 \end{pmatrix}$$

Eigenvalues: $\quad (85-\lambda)(40-\lambda) - 900 = 0 \Rightarrow \lambda_1 = 100,\ \lambda_2 = 25$

Orthonormal eigenbasis: $\quad \bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix} \perp \bar{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$$A\bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 10 \\ -20 \end{pmatrix} \perp A\bar{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 10 \\ 5 \end{pmatrix}$$

# SVD: Example

Consider the unit circle

$$\bar{x} = \bar{v}_1 \cos t + \bar{v}_2 \sin t$$

What is its image under the linear map $L$?

# SVD: Example

Consider the unit circle

$$\bar{x} = \bar{v}_1 \cos t + \bar{v}_2 \sin t$$

What is its image under the linear map $L$?

$$L\bar{x} = A\bar{v}_1 \cos t + A\bar{v}_2 \sin t, \; A\bar{v}_1 \perp A\bar{v}_2$$

# SVD: Example

Consider the unit circle

$$\bar{x} = \bar{v}_1 \cos t + \bar{v}_2 \sin t$$

What is its image under the linear map $L$?

$$L\bar{x} = A\bar{v}_1 \cos t + A\bar{v}_2 \sin t, \; A\bar{v}_1 \perp A\bar{v}_2$$

$\Rightarrow L\bar{x}$ is an ellipse with semi-axes $||A\bar{v}_1||, \; ||A\bar{v}_2||$

# SVD: Example

Consider the unit circle

$$\bar{x} = \bar{v}_1 \cos t + \bar{v}_2 \sin t$$

What is its image under the linear map $L$?

$$L\bar{x} = A\bar{v}_1 \cos t + A\bar{v}_2 \sin t, \; A\bar{v}_1 \perp A\bar{v}_2$$

$$\Rightarrow L\bar{x} \text{ is an ellipse with semi-axes } ||A\bar{v}_1||, \; ||A\bar{v}_2||$$

$$||A\bar{v}_1||^2 = (A\bar{v}_1, A\bar{v}_1) =$$

# SVD: Example

Consider the unit circle

$$\bar{x} = \bar{v}_1 \cos t + \bar{v}_2 \sin t$$

What is its image under the linear map $L$?

$$L\bar{x} = A\bar{v}_1 \cos t + A\bar{v}_2 \sin t,\ A\bar{v}_1 \perp A\bar{v}_2$$

$$\Rightarrow L\bar{x} \text{ is an ellipse with semi-axes } ||A\bar{v}_1||,\ ||A\bar{v}_2||$$

$$||A\bar{v}_1||^2 = (A\bar{v}_1, A\bar{v}_1) = \lambda_1(\bar{v}_1,\ \bar{v}_1) =$$

# SVD: Example

Consider the unit circle

$$\bar{x} = \bar{v}_1 \cos t + \bar{v}_2 \sin t$$

What is its image under the linear map $L$?

$$L\bar{x} = A\bar{v}_1 \cos t + A\bar{v}_2 \sin t,\ A\bar{v}_1 \perp A\bar{v}_2$$

$$\Rightarrow L\bar{x} \text{ is an ellipse with semi-axes } ||A\bar{v}_1||,\ ||A\bar{v}_2||$$

$$||A\bar{v}_1||^2 = (A\bar{v}_1, A\bar{v}_1) = \lambda_1(\bar{v}_1,\ \bar{v}_1) = \lambda_1 \Rightarrow$$

# SVD: Example

Consider the unit circle

$$\bar{x} = \bar{v}_1 \cos t + \bar{v}_2 \sin t$$

What is its image under the linear map $L$?

$$L\bar{x} = A\bar{v}_1 \cos t + A\bar{v}_2 \sin t, \; A\bar{v}_1 \perp A\bar{v}_2$$

$$\Rightarrow L\bar{x} \text{ is an ellipse with semi-axes } ||A\bar{v}_1||, \; ||A\bar{v}_2||$$

$$||A\bar{v}_1||^2 = (A\bar{v}_1, A\bar{v}_1) = \lambda_1(\bar{v}_1, \; \bar{v}_1) = \lambda_1 \Rightarrow ||A\bar{v}_1|| = \sqrt{100} = 10$$

# SVD: Example

Consider the unit circle

$$\bar{x} = \bar{v}_1 \cos t + \bar{v}_2 \sin t$$

What is its image under the linear map $L$?

$$L\bar{x} = A\bar{v}_1 \cos t + A\bar{v}_2 \sin t, \; A\bar{v}_1 \perp A\bar{v}_2$$

$$\Rightarrow L\bar{x} \text{ is an ellipse with semi-axes } ||A\bar{v}_1||, \; ||A\bar{v}_2||$$

$$||A\bar{v}_1||^2 = (A\bar{v}_1, A\bar{v}_1) = \lambda_1(\bar{v}_1, \; \bar{v}_1) = \lambda_1 \Rightarrow ||A\bar{v}_1|| = \sqrt{100} = 10$$

$$||A\bar{v}_2|| = \sqrt{25} = 5$$

# SVD: Example

Consider the unit circle

$$\bar{x} = \bar{v}_1 \cos t + \bar{v}_2 \sin t$$

What is its image under the linear map $L$?

$$L\bar{x} = A\bar{v}_1 \cos t + A\bar{v}_2 \sin t, \; A\bar{v}_1 \perp A\bar{v}_2$$

$$\Rightarrow L\bar{x} \text{ is an ellipse with semi-axes } ||A\bar{v}_1||, \; ||A\bar{v}_2||$$

$$||A\bar{v}_1||^2 = (A\bar{v}_1, A\bar{v}_1) = \lambda_1(\bar{v}_1, \; \bar{v}_1) = \lambda_1 \Rightarrow ||A\bar{v}_1|| = \sqrt{100} = 10$$

$$||A\bar{v}_2|| = \sqrt{25} = 5$$

*The eigenvalues of $A^T A$ define the ellipse as the image of the unit circle.*

# Singular Values

**Definition:** The singular values of a matrix $A_{n\times m}$ are the square roots of the eigenvalues of the symmetric matrix $(A^T A)_{m\times m}$ listed with their algebraic multiplicities:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \geq 0$$

# Singular Values

**Definition:** The singular values of a matrix $A_{n\times m}$ are the square roots of the eigenvalues of the symmetric matrix $(A^TA)_{m\times m}$ listed with their algebraic multiplicities:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \geq 0$$

**Theorem:** Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$, $L\bar{x} = A\bar{x}$ be invertible. The image of the unit circle under the map $L$ is an ellipse $E$. Singular values of $A$ are the length of semi-axes of $E$.

# Example

$$L\colon \mathbb{R}^3 \to \mathbb{R}^2,\ L\bar{x} = A\bar{x} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \bar{x}$$

$$A^T A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

$$(1-\lambda)^2(2-\lambda) - 2 = 0 \Rightarrow \lambda_1 = 3,\ \lambda_2 = 1,\ \lambda_3 = 0$$

$$\sigma_1 = \sqrt{3} > \sigma_2 = 1 > \sigma_3 = 0$$

$$\bar{v}_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \bar{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \bar{v}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

# Example

$$A\bar{v}_1 = \frac{1}{\sqrt{6}}\left(\begin{array}{c} 3 \\ 3 \end{array}\right), A\bar{v}_2 = \frac{1}{\sqrt{2}}\left(\begin{array}{c} -1 \\ 1 \end{array}\right), A\bar{v}_3 = \frac{1}{\sqrt{3}}\left(\begin{array}{c} 0 \\ 0 \end{array}\right)$$

$$||A\bar{v}_1|| = \sqrt{3} = \sigma_1,\ ||A\bar{v}_2|| = 1 = \sigma_2,\ ||A\bar{v}_3|| = 0 = \sigma_3$$

The unit sphere in $\mathbb{R}^3$ is defined by

$$\bar{x} = c_1\bar{v}_1 + c_2\bar{v}_2 + c_3\bar{v}_3, \quad c_1^2 + c_2^2 + c_3^2 = 1$$

The image of the unit sphere is

$$L\bar{x} = c_1 L\bar{v}_1 + c_2 L\bar{v}_2 = c_1\lambda_1\bar{v}_1 + c_2\lambda_2\bar{v}_2$$

$$c_1^2 + c_2^2 \leq 1$$

an ellipse

# Singular Value Decomposition

**Lemma:** If $rank\, A_{n\times m} = r$, then its singular values

$$\sigma_1, \dots, \sigma_r \neq 0 \quad \text{and} \quad \sigma_{r+1}, \dots, \sigma_m = 0$$

**Theorem (SVD):** Any matrix $A_{n\times m}$ can be represented in the form

$$A = U\Sigma V^T,$$

$U$ is an orthogonal $n \times n$ matrix

$V$ is an orthogonal $m \times m$ matrix

$\Sigma$ is an $n \times n$ matrix whose first $r$ diagonal entries are nonzero singular values of $A$, $r = rank\, A$, and all other entries vanish

# Singular Value Decomposition: Remarks

**Remark 1:**

$$A\bar{v}_i = \sigma_i \bar{u}_i,\ i = 1, \dots, r \qquad A\bar{v}_i = \bar{0},\ i = r+1, \dots, m$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$Im\, A = span(\bar{u}_1, \dots, \bar{u}_r) \qquad Ker\, A = span(\bar{v}_{r+1}, \dots, \bar{v}_m)$$

**Remark 2:** $A = U\Sigma V^T \Rightarrow A^T = V\Sigma \underbrace{U^T}_{U^{-1}} \Rightarrow A^T U = V\Sigma^T$

$$A^T\bar{u}_i = \sigma_i \bar{v}_i,\ i = 1, \dots, r \qquad A^T\bar{u}_i = \bar{0},\ i = r+1, \dots, m$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$Im\, A^T = span(\bar{v}_1, \dots, \bar{v}_r) \qquad Ker\, A^T = span(\bar{u}_{r+1}, \dots, \bar{u}_m)$$

# Singular Value Decomposition: Example 1

$$A = \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix}$$

$$\bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \; \bar{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow V = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$$

$$\bar{u}_1 = \frac{1}{\sigma_1} A\bar{v}_1 = \frac{1}{10\sqrt{5}} \begin{pmatrix} 10 \\ -20 \end{pmatrix}, \; \bar{u}_2 = \frac{1}{\sigma_2} A\bar{v}_2 = \frac{1}{5\sqrt{5}} \begin{pmatrix} 10 \\ 5 \end{pmatrix}$$

$$\Rightarrow U = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix}$$

$$A = \begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

# Singular Value Decomposition: Example 2

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$\bar{v}_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \bar{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \bar{v}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$\bar{u}_1 = \frac{1}{\sigma_1} A\bar{v}_1 = \frac{1}{\sqrt{3}\sqrt{6}} \begin{pmatrix} 3 \\ 3 \end{pmatrix}, \ \bar{u}_2 = \frac{1}{\sigma_2} A\bar{v}_2 = \frac{1}{1 \cdot \sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\Rightarrow U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}$$