# Lecture 3
# Eigenvalue, Matrix Decomposition

- Motivation
- Eigenvalues and Eigenvectors
- Properties about Eigenvalues
- Eigenvalue Decomposition
- Singular Value Decomposition

# Contents

# Example

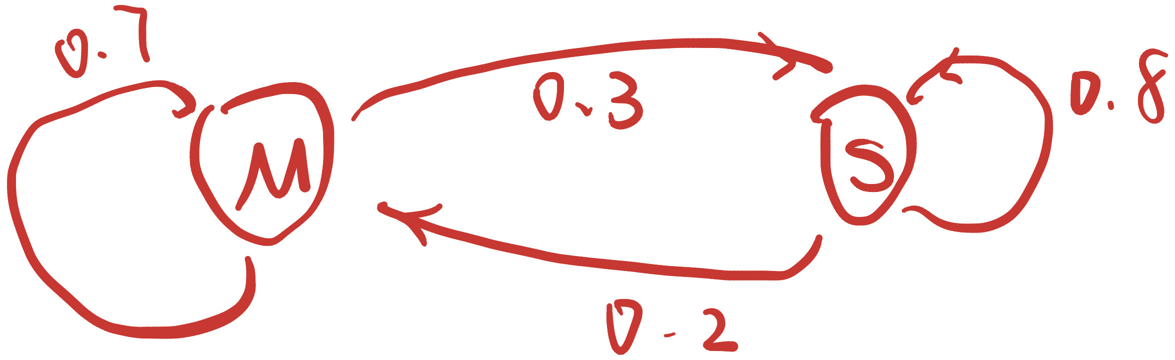
- In a certain town, $30\%$ married men get divorced each year and $20\%$ single men get married each year.
- There are 8000 married men and 2000 single men at the beginning. Assume the total population always remains constant.
- Let $\mathbf{w}_0 = \begin{pmatrix} 8000 \\ 2000 \end{pmatrix}$ be the initial status, and $\mathbf{w}_i$ denote the status after $i$ years.
- What is the marital status when time goes to infinity? How about we change $\mathbf{w}_0$?

# Example

- In a certain town, $30\%$ married men get divorced each year and $20\%$ single men get married each year.
- There are 8000 married men and 2000 single men at the beginning. Assume the total population always remains constant.
- Let $\mathbf{w}_0 = \begin{pmatrix} 8000 \\ 2000 \end{pmatrix}$ be the initial status, and $\mathbf{w}_i$ denote the status after $i$ years.
- What is the marital status when time goes to infinity? How about we change $\mathbf{w}_0$?

# Example

- In a certain town, $30\%$ married men get divorced each year and $20\%$ single men get married each year.
- There are 8000 married men and 2000 single men at the beginning. Assume the total population always remains constant.
- Let $\mathbf{w}_0 = \begin{pmatrix} 8000 \\ 2000 \end{pmatrix}$ be the initial status, and $\mathbf{w}_i$ denote the status after $i$ years.
- What is the marital status when time goes to infinity? How about we change $\mathbf{w}_0$?




# Example

$w^* = \begin{pmatrix} 4000 \\ 6000 \end{pmatrix}$ $\quad Aw^* = w^*$

- In a certain town, $30\%$ married men get divorced each year and $20\%$ single men get married each year.
- There are 8000 married men and 2000 single men at the beginning. Assume the total population always remains constant.
- Let $\mathbf{w}_0 = \begin{pmatrix} 8000 \\ 2000 \end{pmatrix}$ be the initial status, and $\mathbf{w}_i$ denote the status after $i$ years.

$$W_i = \begin{pmatrix} \text{Married after i-th year} \\ \text{Single after i-th year} \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix} W_{i-1}$$

- What is the marital status when time goes to infinity? How about we change $\mathbf{w}_0$?

$0.7\, W_{i-1,1} + 0.2\, W_{i-1,2}$

$0.3\, W_{i-1,1} + 0.8\, W_{i-1,2}$

$$W_t = A W_{t-1} = A^2 W_{t-2} = \cdots = A^t W_0$$

# Example

- Let

$$A = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}.$$

  and $\mathbf{w}_i = A\mathbf{w}_{i-1} = A^i\mathbf{w}_0$.

- Using computer, we find $\mathbf{w}_n \to \begin{pmatrix} 4000 \\ 6000 \end{pmatrix}$ as $n \to \infty$.

- We represent a general initial marital status as

$$\mathbf{w}_0 = x_1\mathbf{u}_1 + x_2\mathbf{u}_2, \text{ where } \mathbf{u}_1 = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

  We check that $A\mathbf{u}_1 = \mathbf{u}_1$ and $A\mathbf{u}_2 = 0.5\mathbf{u}_2$

- Then the marital status after $n$ years is

$$A^n\mathbf{w}_0 = x_1A^n\mathbf{u}_1 + x_2A^n\mathbf{u}_2 = x_1\mathbf{u}_1 + x_20.5^n\mathbf{u}_2 \to x_1\mathbf{u}_1.$$

# Example

- Let
$$A = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}.$$
and $\mathbf{w}_i = A\mathbf{w}_{i-1} = A^i\mathbf{w}_0$.
- Using computer, we find $\mathbf{w}_n \to \begin{pmatrix} 4000 \\ 6000 \end{pmatrix}$ as $n \to \infty$.
- We represent a general initial marital status as
$$\mathbf{w}_0 = x_1\mathbf{u}_1 + x_2\mathbf{u}_2, \text{ where } \mathbf{u}_1 = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$
We check that $A\mathbf{u}_1 = \mathbf{u}_1$ and $A\mathbf{u}_2 = 0.5\mathbf{u}_2$
- Then the marital status after $n$ years is
$$A^n\mathbf{w}_0 = x_1A^n\mathbf{u}_1 + x_2A^n\mathbf{u}_2 = x_1\mathbf{u}_1 + x_2 0.5^n\mathbf{u}_2 \to x_1\mathbf{u}_1.$$

# Example

$0.4x_1 + x_2 = 8000$

$0.6x_1 - x_2 = 2000$

- Let

$\Rightarrow x_1 = 10000$

$$A = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}.$$

$Au_1 = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}\begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix} = \begin{pmatrix} 0.28+0.12 \\ 0.12+0.48 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$

and $\mathbf{w}_i = A\mathbf{w}_{i-1} = A^i\mathbf{w}_0$.

- Using computer, we find $\mathbf{w}_n \to \begin{pmatrix} 4000 \\ 6000 \end{pmatrix}$ as $n \to \infty$.

$Au_2 = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0.7-0.2 \\ 0.3-0.8 \end{pmatrix} = \begin{pmatrix} +0.5 \\ -0.5 \end{pmatrix}$

- We represent a general initial marital status as

$$\mathbf{w}_0 = x_1\mathbf{u}_1 + x_2\mathbf{u}_2, \text{ where } \mathbf{u}_1 = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$A^n u_1 = A^{n-1}(Au_1) = A^{n-1}u_1$

We check that $A\mathbf{u}_1 = \mathbf{u}_1$ and $A\mathbf{u}_2 = 0.5\mathbf{u}_2$

$A^n u_2 = A^{n-1}(Au_2) = 0.5A^{n-1}u_2$

- Then the marital status after $n$ years is

$A^n(x_1u_1 + x_2u_2)$

$$A^n\mathbf{w}_0 = x_1A^n\mathbf{u}_1 + x_2A^n\mathbf{u}_2 = x_1\mathbf{u}_1 + x_2 0.5^n\mathbf{u}_2 \to x_1\mathbf{u}_1.$$

# Contents

# Definition

- Let $A$ be an $n \times n$ matrix.
- A scalar $\lambda$ is said to be an **eigenvalue** of $A$ if there exists a nonzero vector $\mathbf{x}$ such that

eigenvalue

$$A\mathbf{x} = \lambda\mathbf{x}.$$

eigenvector associated with $\lambda$

- The vector $\mathbf{x}$ is said to be an **eigenvector belonging to** $\lambda$.

$(\lambda, x)$ is an eigen-pair of $A$

# Implications

$$Ax = \lambda x$$

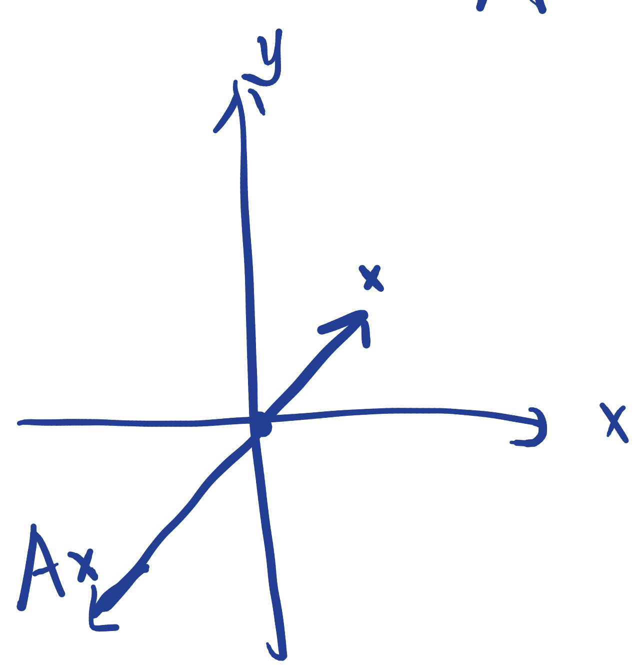
$$cAx = \lambda x \qquad c \neq 0 \iff A(cx) = \lambda(cx)$$

- An eigenvector $\mathbf{x}$ and $A\mathbf{x}$ have the same direction.
- If $\mathbf{x}$ is an eigenvector belonging to $\lambda$, so is $c\mathbf{x}$ for any $c \neq 0$.
- If $\mathbf{x}$ is an eigenvector of $A$ belonging to $\lambda$, then $\mathbf{x}$ is an eigenvector of $A^s$ belonging to $\lambda^s$.

$$Ax = \lambda x \Rightarrow A^s x = A^{s-1}(Ax) = \lambda A^{s-1} x = \cdots = \lambda^s x$$

$s \in \mathbb{N}_+$

$$x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$




$$A = \begin{pmatrix} 0 & -3 \\ 1 & -4 \end{pmatrix} \qquad Ax = \begin{pmatrix} -3 \\ -3 \end{pmatrix}$$

# Example

- Let

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

- Then

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

- Thus, $-3$ is the eigenvalue of $A$ and the corresponding eigenvector is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

# Equivalent Characterizations

- Note that $A\mathbf{x} = \lambda\mathbf{x}$ is equivalent to

$\boxed{A}\,\boxed{x} = \lambda\,\boxed{x}$

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

$I = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$

Note that it is not $(A - \lambda)\mathbf{x} = \mathbf{0}$.

- The following statements are equivalent:
  - $\lambda$ is an eigenvalue of $A$.
  - $(A - \lambda I)\mathbf{x} = 0$ has a nontrivial solution.
  - $\mathcal{N}(A - \lambda I) \neq \{\mathbf{0}\}$.
  - $A - \lambda I$ is singular.
  - $\det(A - \lambda I) = 0$.
- $\mathcal{N}(A - \lambda I)$ is called the **eigenspace** of eigenvalue $\lambda$.
- All nonzero vectors in $\mathcal{N}(A - \lambda I)$ are eigenvectors corresponding to $\lambda$.

# Equivalent Characterizations

- Note that $A\mathbf{x} = \lambda\mathbf{x}$ is equivalent to

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

  Note that it is not $(A - \lambda)\mathbf{x} = \mathbf{0}$.

- The following statements are equivalent:
  - $\lambda$ is an eigenvalue of $A$.
  - $(A - \lambda I)\mathbf{x} = 0$ has a nontrivial solution.
  - $\mathcal{N}(A - \lambda I) \neq \{\mathbf{0}\}$.
  - $A - \lambda I$ is singular.
  - $\det(A - \lambda I) = 0$.

- $\mathcal{N}(A - \lambda I)$ is called the **eigenspace** of eigenvalue $\lambda$.

- All nonzero vectors in $\mathcal{N}(A - \lambda I)$ are eigenvectors corresponding to $\lambda$.

# Equivalent Characterizations

- Note that $A\mathbf{x} = \lambda\mathbf{x}$ is equivalent to

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

  Note that it is not $(A - \lambda)\mathbf{x} = \mathbf{0}$.

- The following statements are equivalent:
  - $\lambda$ is an eigenvalue of $A$.
  - $(A - \lambda I)\mathbf{x} = 0$ has a nontrivial solution.
  - $\mathcal{N}(A - \lambda I) \neq \{\mathbf{0}\}$.
  - $A - \lambda I$ is singular.
  - $\det(A - \lambda I) = 0$.
- $\mathcal{N}(A - \lambda I)$ is called the **eigenspace** of eigenvalue $\lambda$.
- All nonzero vectors in $\mathcal{N}(A - \lambda I)$ are eigenvectors corresponding to $\lambda$.

# Equivalent Characterizations

- Note that $A\mathbf{x} = \lambda\mathbf{x}$ is equivalent to

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

  Note that it is not $(A - \lambda)\mathbf{x} = \mathbf{0}$.

- The following statements are equivalent:
  - $\lambda$ is an eigenvalue of $A$.
  - $(A - \lambda I)\mathbf{x} = 0$ has a nontrivial solution.
  - $\mathcal{N}(A - \lambda I) \neq \{\mathbf{0}\}$.
  - $A - \lambda I$ is singular.
  - $\det(A - \lambda I) = 0$.
- $\mathcal{N}(A - \lambda I)$ is called the **eigenspace** of eigenvalue $\lambda$.
- All nonzero vectors in $\mathcal{N}(A - \lambda I)$ are eigenvectors corresponding to $\lambda$.

# Characteristic polynomial

• Step 2: Solve chara. equation.

Find $\lambda$ s.t. $p(\lambda)=0$

$\lambda_1 = 4 \qquad \lambda_2 = -3$

- $p(\lambda) = \det(A - \lambda I)$ is an $n$th degree polynomial in $\lambda$.
- $p(\lambda)$ is called the **characteristic polynomial** of $A$.
- $p(\lambda) = 0$ is called the **characteristic equation** of $A$.
- A scalar $\lambda$ is an eigenvalue of $A$ if and only if $p(\lambda) = 0$.

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

• Step 1: write down characteristic polynomial.

$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 3-\lambda & 2 \\ 3 & -2-\lambda \end{vmatrix} = (3-\lambda)(-2-\lambda) - 6$

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

# Example

- The characteristic polynomial of $A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$ is

$$p(\lambda) = \begin{vmatrix} 3-\lambda & 2 \\ 3 & -2-\lambda \end{vmatrix} = \lambda^2 - \lambda - 12 = (\lambda - 4)(\lambda + 3) = 0.$$

- Hence, the the eigenvalues of $A$ are $\lambda_1 = 4$ and $\lambda_2 = -3$.

- The eigenvectors belonging to $\lambda_1$ are nonzero solutions of

$\begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$

$$(A - 4I)\mathbf{x} = \mathbf{0}.$$

$\begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix}$

$c\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, where $c \neq 0$

$\Rightarrow x_1 = 2x_2$

The eigenvectors belonging to $\lambda_2$ are nonzero solutions of

$\begin{pmatrix} 6 & 2 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$

$$(A + 3I)\mathbf{x} = \mathbf{0}.$$

$\begin{pmatrix} 6 & 2 \\ 3 & 1 \end{pmatrix}$

$c \cdot \begin{pmatrix} 1 \\ -3 \end{pmatrix}$, where $c \neq 0$

$\lambda_1 = 0$ $\lambda_2 = 1$

# Example

$$(A-\lambda_1 I)x = 0 = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$\left.\begin{aligned} 2x_1 - 3x_2 + x_3 &= 0 \\ x_1 - 2x_2 + x_3 &= 0 \\ x_1 - 3x_2 + 2x_3 &= 0 \end{aligned}\right\} \Rightarrow x_1 = x_2 = x_3$$

$$\Rightarrow c\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad c \neq 0$$

Find the eigenvalues and the corresponding eigenvectors of

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

① $A - \lambda I = \begin{pmatrix} 2-\lambda & -3 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & -3 & 2-\lambda \end{pmatrix}$

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

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

$$(A-\lambda_2 I)x = 0$$

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

$$\Rightarrow x_1 - 3x_2 + x_3 = 0.$$

$$\left\{ \begin{pmatrix} 3x_2 - x_3 \\ x_2 \\ x_3 \end{pmatrix} : x_2, x_3 \in \mathbb{R} \right\}$$

# Example

$i = \sqrt{-1}$

Find the eigenvalues and the corresponding eigenvectors of

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

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

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

$$\lambda - 1 = \pm\sqrt{-4} = 2i$$

$$\lambda_1 = 1 + 2i \qquad \lambda_2 = 1 - 2i$$

# Complex Eigenvalues of Real Matrices

- As $p(\lambda)$ has degree $n$, $p(\lambda)$ can be factored into the product of $n$ linear terms:

$$p(\lambda) = c(\lambda_1 - \lambda)\cdots(\lambda_n - \lambda), \tag{1}$$

  where $\lambda_i$ is the root of $p(\lambda)$.

- For real valued matrices,
  - Complex eigenvalues occur in **conjugate pairs**, i.e., if $\lambda$ is an eigenvalue, so is $\bar{\lambda}$.
  - If $\mathbf{z}$ is an eigenvector belonging to a complex eigenvalue $\lambda$, then $\bar{\mathbf{z}}$ is an eigenvector belonging to $\bar{\lambda}$.

# Complex Eigenvalues of Real Matrices

- As $p(\lambda)$ has degree $n$, $p(\lambda)$ can be factored into the product of $n$ linear terms:

$$p(\lambda) = c(\lambda_1 - \lambda) \cdots (\lambda_n - \lambda), \tag{1}$$

  where $\lambda_i$ is the root of $p(\lambda)$.

$Ax = \lambda x \Rightarrow \overline{Ax} = \overline{\lambda x}$

$\bar{A}\bar{x} = \bar{\lambda}\bar{x}$

- For real valued matrices,
  - Complex eigenvalues occur in **conjugate pairs**, i.e., if $\lambda$ is an eigenvalue, so is $\bar{\lambda}$.

$\lambda = a + bi \quad \bar{\lambda} = a - bi$

$\Rightarrow A\bar{x} = \bar{\lambda}\bar{x}$

  - If $\mathbf{z}$ is an eigenvector belonging to a complex eigenvalue $\lambda$, then $\bar{\mathbf{z}}$ is an eigenvector belonging to $\bar{\lambda}$.

$(\lambda, x)$ is eigen-pair

$\Rightarrow (\bar{\lambda}, \bar{x})$ is another eigen-pair

# Contents

# Multiplicity of Eigenvalues

- We know that $\lambda_1, \ldots, \lambda_n$ may not be all distinct.
- Let $\lambda_1, \ldots, \lambda_p$ be the $p$ distinct eigenvalues.
- The eigenvalue $\lambda_k$ has *multiplicity* $m_k$. We know that $\sum_k m_k = n$.
- The characteristic polynomial can be written as

$$p(\lambda) = c(\lambda_1 - \lambda)^{m_1} \cdots (\lambda_p - \lambda)^{m_p}. \tag{2}$$

- Example: For $p(\lambda) = (1-\lambda)^2(4-\lambda)^3$, the multiplicity of $1$ is $2$ and the multiplicity of $4$ is $3$.

# Multiplicity of Eigenvalues

- We know that $\lambda_1, \ldots, \lambda_n$ may not be all distinct.
- Let $\lambda_1, \ldots, \lambda_p$ be the $p$ distinct eigenvalues.
- The eigenvalue $\lambda_k$ has *multiplicity* $m_k$. We know that $\sum_k m_k = n$.
- The characteristic polynomial can be written as

$$p(\lambda) = c(\lambda_1 - \lambda)^{m_1} \cdots (\lambda_p - \lambda)^{m_p}. \tag{2}$$

- Example: For $p(\lambda) = (1-\lambda)^2(4-\lambda)^3$, the multiplicity of $1$ is $2$ and the multiplicity of $4$ is $3$.

# Multiplicity of Eigenvalues

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

- We know that $\lambda_1, \ldots, \lambda_n$ may not be all distinct.
- Let $\lambda_1, \ldots, \lambda_p$ be the $p$ distinct eigenvalues.
- The eigenvalue $\lambda_k$ has *multiplicity* $m_k$. We know that $\sum_k m_k = n$.
- The characteristic polynomial can be written as

$$p(\lambda) = c(\lambda_1 - \lambda)^{m_1} \cdots (\lambda_p - \lambda)^{m_p}. \tag{2}$$

- Example: For $p(\lambda) = (1-\lambda)^2(4-\lambda)^3$, the multiplicity of $1$ is $2$ and the multiplicity of $4$ is $3$.

# Multiplicity of Eigenvalues

- We know that $\lambda_1, \ldots, \lambda_n$ may not be all distinct.
- Let $\lambda_1, \ldots, \lambda_p$ be the $p$ distinct eigenvalues.
- The eigenvalue $\lambda_k$ has *multiplicity* $m_k$. We know that $\sum_k m_k = n$.
- The characteristic polynomial can be written as

$$p(\lambda) = c(\lambda_1 - \lambda)^{m_1} \cdots (\lambda_p - \lambda)^{m_p}. \tag{2}$$

- Example: For $p(\lambda) = (1-\lambda)^2(4-\lambda)^3$, the multiplicity of $1$ is $2$ and the multiplicity of $4$ is $3$.

# Product and Sum of the Eigenvalues

- Consider an $n \times n$ square matrix $A = (a_{ij})$.
- Let $\lambda_1, \ldots, \lambda_n$ be the $n$ eigenvalues of $A$.
- $\prod_{i=1}^{n} \lambda_i = \det(A)$.
- $\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}$.
- Proof:

# Product and Sum of the Eigenvalues

① degree $n$ term: $c(-\lambda)^n = (-\lambda)^n \Rightarrow c=1$

③ $p(\lambda) = \det(A-\lambda I)$

$p(0) = \det(A) = c\,\lambda_1 \cdots \lambda_n = \prod_{i}^{n} \lambda_i$

② degree $(n-1)$-term: $c\sum \lambda_i (-\lambda)^{n-1} = \sum a_{ii} (-\lambda)^{n-1} \Rightarrow c\sum \lambda_i = \sum a_{ii} \Rightarrow \sum \lambda_i = \mathrm{Tr}(A)$

- Consider an $n \times n$ square matrix $A = (a_{ij})$.
- Let $\lambda_1, \ldots, \lambda_n$ be the $n$ eigenvalues of $A$.
- $\prod_{i=1}^{n} \lambda_i = \det(A)$.
- $\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}$.
- Proof:

$p(\lambda) = c(\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$

$= \det(A - \lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}$

$(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$ accounts for degree $n$ and degree $(n-1)$ terms

# Transpose and Inverse

- As $|A - \lambda I| = |A^\top - \lambda I|$, $A$ and $A^\top$ have same characteristic polynomial, and hence the same eigenvalues.
- If $A$ is singular, $0$ is an eigenvalue of $A$.
- If $A$ is invertible, $\lambda$ is an eigenvalue of $A$ if and only if $\lambda^{-1}$ is an eigenvalue of $A^{-1}$.
- Proof:

# Transpose and Inverse

- As $|A - \lambda I| = |A^\top - \lambda I|$, $A$ and $A^\top$ have same characteristic polynomial, and hence the same eigenvalues.
- If $A$ is singular, $0$ is an eigenvalue of $A$. $Ax = 0 = 0 \cdot x$
- If $A$ is invertible, $\lambda$ is an eigenvalue of $A$ if and only if $\lambda^{-1}$ is an eigenvalue of $A^{-1}$.
- Proof: $Ax = \lambda x$

$$x = A^{-1}(Ax) = A^{-1}(\lambda x) = \lambda A^{-1} x$$

$$\Rightarrow A^{-1} x = \lambda^{-1} x$$

# Stochastic Matrix

$A = \begin{pmatrix} 0.3 & 0.2 \\ 0.7 & 0.8 \end{pmatrix}$

- An $n \times n$ matrix $A$ is a stochastic matrix if
  1. all the entries are non-negative ($a_{ij} \geq 0$);
  2. the summation of each column is $1$ ($\mathbf{1}^\top A = \mathbf{1}^\top$).

$A = [a_1, \ldots, a_n]$

$\mathbb{1}^T A = [\mathbb{1}^T a_1, \cdots, \mathbb{1}^T a_n]$

$= [1, \cdots, 1]$

$= \mathbb{1}^T$

- For any vector $\mathbf{x} \in \mathbb{R}^n$, $A\mathbf{x}$ and $\mathbf{x}$ have the same sum.
- $1$ is an eigenvalue of $A$ (and $A^\top$).
- All the eigenvalues $\lambda$ of $A$ have $|\lambda| \leq 1$.
- Proof:

$\mathbb{1}^T (Ax) = (\mathbb{1}^T A) x = \mathbb{1}^T x$

$A^T \mathbb{1} = \mathbb{1} \Rightarrow A^T \mathbb{1} = (1) \cdot \mathbb{1}$

$A^n x = \lambda^n x$

LHS absolute values are bounded

# Contents

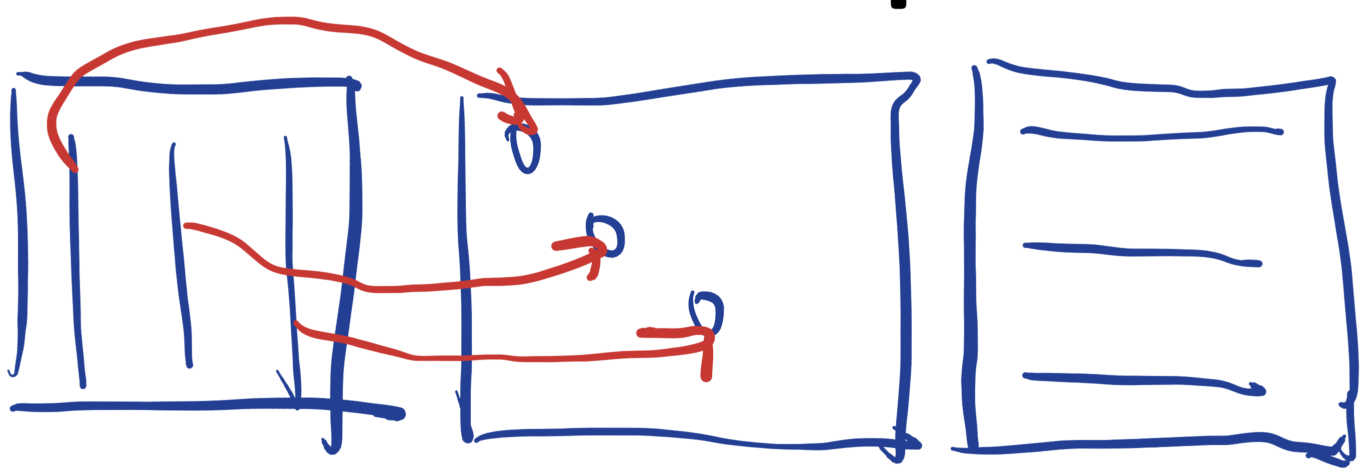
# Spectral Theorem

- If $A$ is a real symmetric matrix, the spectral theorem shows that there exists an orthogonal matrix that diagonalize $A$.
- Every real symmetric matrix $A$ can be factored into $Q\Lambda Q^\top$ where $Q$ is an orthogonal matrix and $\Lambda$ is a real diagonal matrix.
-
  - The diagonal entries $\lambda_i$ of $\Lambda$ are eigenvalues of $A$.
  - The columns $\mathbf{q}_i$ of $Q$ are eigenvectors belonging to $\lambda_i$, respectively.

# Spectral Theorem

- If $A$ is a real symmetric matrix, the spectral theorem shows that there exists an orthogonal matrix that diagonalize $A$.
- Every real symmetric matrix $A$ can be factored into $Q\Lambda Q^\top$ where $Q$ is an orthogonal matrix and $\Lambda$ is a real diagonal matrix.
- 
  - The diagonal entries $\lambda_i$ of $\Lambda$ are eigenvalues of $A$.
  - The columns $\mathbf{q}_i$ of $Q$ are eigenvectors belonging to $\lambda_i$, respectively.

# Spectral Theorem

$A =$ 

- If $A$ is a real symmetric matrix, the spectral theorem shows that there exists an orthogonal matrix that diagonalize $A$.
- Every real symmetric matrix $A$ can be factored into $Q\Lambda Q^\top$ where $Q$ is an orthogonal matrix and $\Lambda$ is a real diagonal matrix.
-
  - The diagonal entries $\lambda_i$ of $\Lambda$ are eigenvalues of $A$.
  - The columns $\mathbf{q}_i$ of $Q$ are eigenvectors belonging to $\lambda_i$, respectively.

Orthogonal matrix: $Q \in R^{n\times n}$ $\quad Q^T Q = I_n \quad = (q_i^T q_j)_{i,j}$

$Q = [q_1, \dots, q_n]$ $\quad Q^T Q = \begin{pmatrix} q_1^T \\ \vdots \\ q_n^T \end{pmatrix} (q_1, \dots, q_n) = \begin{pmatrix} q_1^T q_1 & q_1^T q_2 & \cdots & q_1^T q_n \\ & \ddots & & \\ & & \ddots & \\ q_n^T q_1 & \cdots & & q_n^T q_n \end{pmatrix}$

# Spectral Theorem

We can also write

$$A = \begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^\top \\ \vdots \\ \mathbf{q}_n^\top \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 \mathbf{q}_1^\top \\ \vdots \\ \lambda_n \mathbf{q}_n^\top \end{bmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i \mathbf{q}_i \mathbf{q}_i^\top$$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$

$\mathrm{Tr}(q_i q_i^T) = \mathrm{Tr}(q_i^T q_i) = \|q_i\|_2^2 = 1$

# Properties of $A^\top A$

Let $A$ be any $m \times n$ matrix $A$ of real numbers.

- $A^\top A$ is symmetric.
- $A^\top A$ is diagonalizable by an orthogonal matrix, and the eigenvalues of $A^\top A$ are real.
- $\mathcal{N}(A^\top A) = \mathcal{N}(A)$.
- $\mathrm{rank}(A) = \mathrm{rank}(A^\top A)$.
- The eigenvalues of $A^\top A$ are nonnegative.

# Contents

# Singular Value Decomposition

The **singular-value decomposition (SVD)** of an $m \times n$ matrix $A$ of real numbers is a factorization of the form $U\Sigma V^\top$, where

- $U$ is an $m \times m$ orthogonal matrix;
- $V$ is an $n \times n$ orthogonal matrix;
- $\Sigma$ is an $m \times n$ matrix whose off-diagonal entries are all $0$'s, and whose diagonal entries $\sigma_i$, $i = 1, \ldots, n$, called the **singular values** satisfy

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

Here singular values $\sigma_k = 0$ if $k > \min\{m, n\}$.

SVD exists for any real matrix.

# Linear Transformation View

- Suppose $A$ has the SVD: $A = U\Sigma V^\top$.
- The columns of $U$ form an orthonormal basis of $\mathbb{R}^m$.
- The columns of $V$ form an orthonormal basis of $\mathbb{R}^n$.
- The linear transformation $L(\mathbf{x}) = A\mathbf{x}$ has the matrix representation $\Sigma$ with respect to the above bases of $\mathbb{R}^m$ and $\mathbb{R}^n$.
- In other words,

$$L(\mathbf{x}) = U\Sigma V^\top \mathbf{x}$$

$$[L(\mathbf{x})]_U = \Sigma[\mathbf{x}]_V$$

$$L(\mathbf{x}) = U$$

$$[\mathbf{x}]_V = V^\top \mathbf{x}$$

# Linear Transformation View

- Suppose $A$ has the SVD: $A = U\Sigma V^\top$.
- The columns of $U$ form an orthonormal basis of $\mathbb{R}^m$.
- The columns of $V$ form an orthonormal basis of $\mathbb{R}^n$.
- The linear transformation $L(\mathbf{x}) = A\mathbf{x}$ has the matrix representation $\Sigma$ with respect to the above bases of $\mathbb{R}^m$ and $\mathbb{R}^n$.
- In other words,

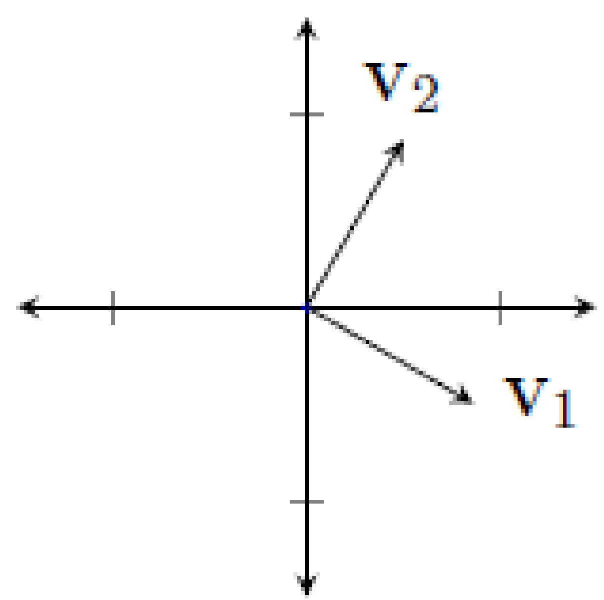
$$L(\mathbf{x}) = U\Sigma V^\top \mathbf{x}$$

$$[L(\mathbf{x})]_U = \Sigma[\mathbf{x}]_V$$

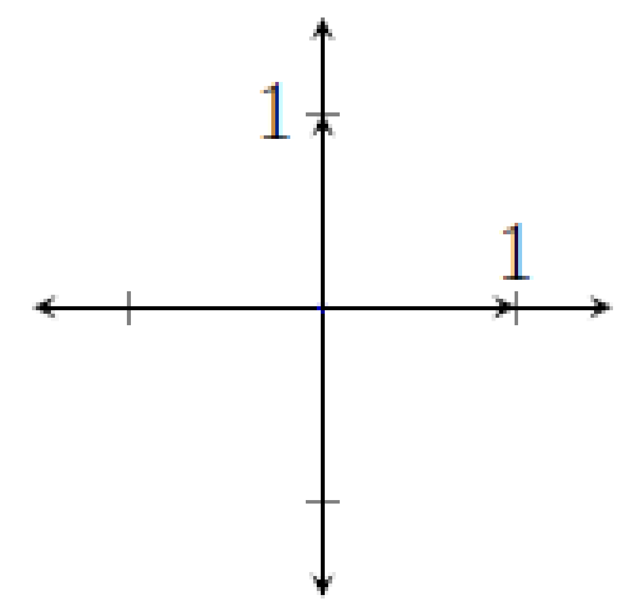
$$L(\mathbf{x}) = U$$

$$[\mathbf{x}]_V = V^\top \mathbf{x}$$

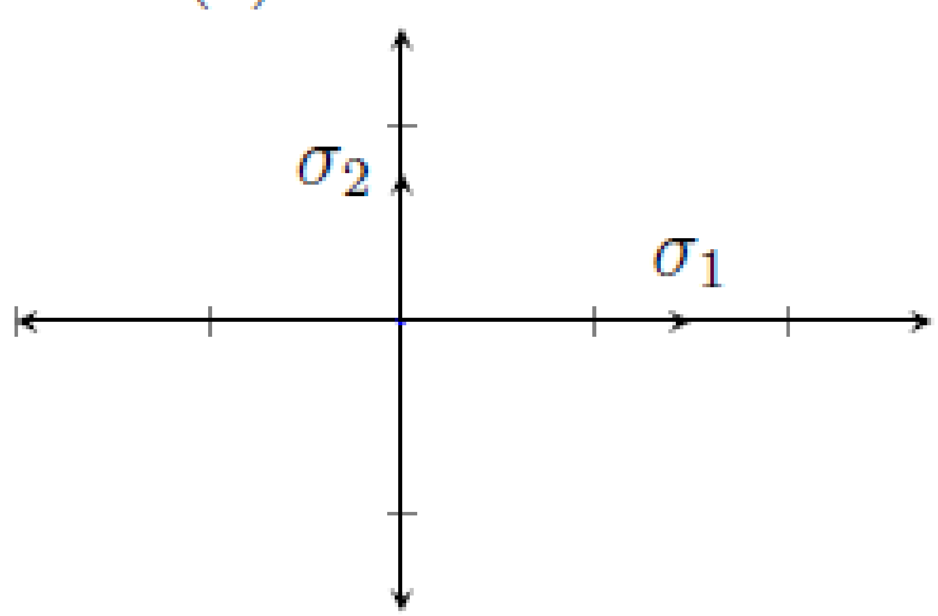
# Visualization of SVD



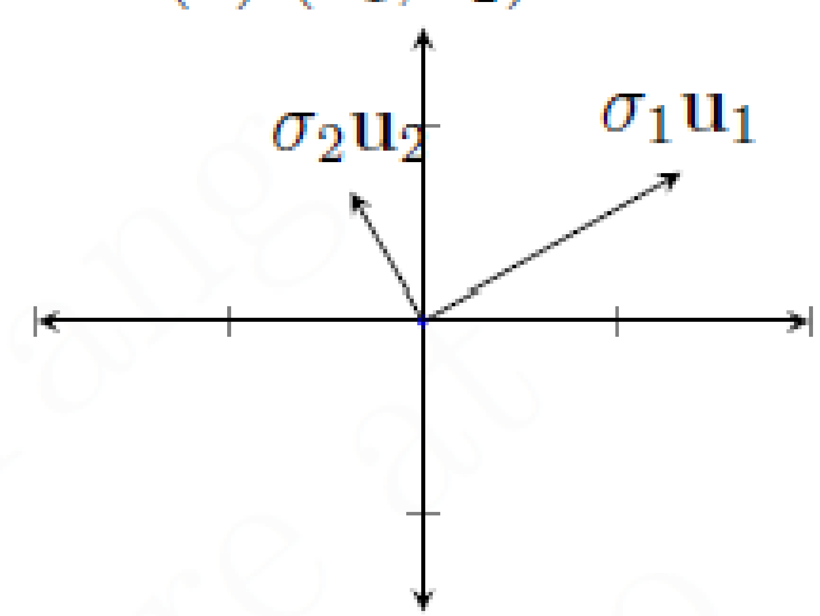

(a) standard axis




(b) $(\mathbf{v}_1, \mathbf{v}_2)$ axis




(c) scaled by $\sigma_1$ and $\sigma_2$, $(\mathbf{u}_1, \mathbf{u}_2)$ axis




(d) standard axis

# Example

1.

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

2.

$$\begin{bmatrix} 2 & 5 & 4 \\ 6 & 3 & 0 \\ 6 & 3 & 0 \\ 2 & 5 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

# Example

1.

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

2.

$$\begin{bmatrix} 2 & 5 & 4 \\ 6 & 3 & 0 \\ 6 & 3 & 0 \\ 2 & 5 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

# SVD and Rank

- If $A = U\Sigma V^\top$, then the rank of $A$ is equal to the number of **nonzero singular values**.
- Proof Technique:
  - Let $r$ be the number of nonzero singular values of $A$.
  - Let $U_r$ and $V_r$ be the first $r$ columns of $U$ and $V$, respectively.
  - Let $\Sigma_r = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$.

  We have $U_r^\top U_r = V_r^\top V_r = I_r$ and

$$A = U_r \Sigma_r V_r^\top. \tag{3}$$

# Outer Product Expansion

$$
\begin{aligned}
A &= U_r \Sigma_r V_r^T \\
&= \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^\top \\ \vdots \\ \mathbf{v}_r^\top \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 \mathbf{v}_1^\top \\ \vdots \\ \sigma_r \mathbf{v}_r^\top \end{bmatrix} \\
&= \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top
\end{aligned}
$$

# Eigenvalues *vs* Singular Values

- For an $n \times n$ square matrix $A$ with SVD and rank $r$

$$A\mathbf{v}_i = \sigma_i \mathbf{u}_i, i = 1, \ldots, r.$$

- The rank of a matrix is always the same as the number of non-zero singular values.

# Eigenvalues *vs* Singular Values

- If $A$ is diagonalizable, i.e., $A$ has $n$ linearly independent eigenvectors $\mathbf{x}_i$

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i, i = 1, \ldots, n$$

  then the rank of the matrix is equal to the number of non-zero eigenvalues.

- But this may not be the case when the matrix is not diagonalizable.

  Consider $M = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

# Property of SVD

Let an $m \times n$ matrix $A$ having SVD $U\Sigma V^\top$. Then

- $\sigma_i = \sqrt{\lambda_i}$, $i = 1, \ldots, n$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ are eigenvalues of $A^\top A$.
- $V$ diagonalizes $A^\top A$, and hence $\mathbf{v}_j$'s are eigenvectors of $A^\top A$.
- the columns of $U$ satisfy:

$$\mathbf{u}_j = \frac{1}{\sigma_j} A\mathbf{v}_j, \quad j = 1, \ldots, r = \operatorname{rank}(A)$$

$$A^\top \mathbf{u}_j = \mathbf{0}, \quad j = r+1, \ldots, m.$$

# Matrix Norm

- Matrix norm $\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$, which is also called *Frobenius* norm.
- If $A = U\Sigma V^\top$, then $\|A\|_F^2 = \sigma_1^2 + \cdots + \sigma_n^2$.

## Lemma

*If $Q$ is an orthogonal matrix, then $\|QA\|_F = \|A\|_F$.*

# Low rank approximation

For a fixed $m \times n$ matrix $A$ and an integer $k$, solve

$$\min_{\text{rank}(S) \le k} \|A - S\|_F. \tag{4}$$

## Theorem

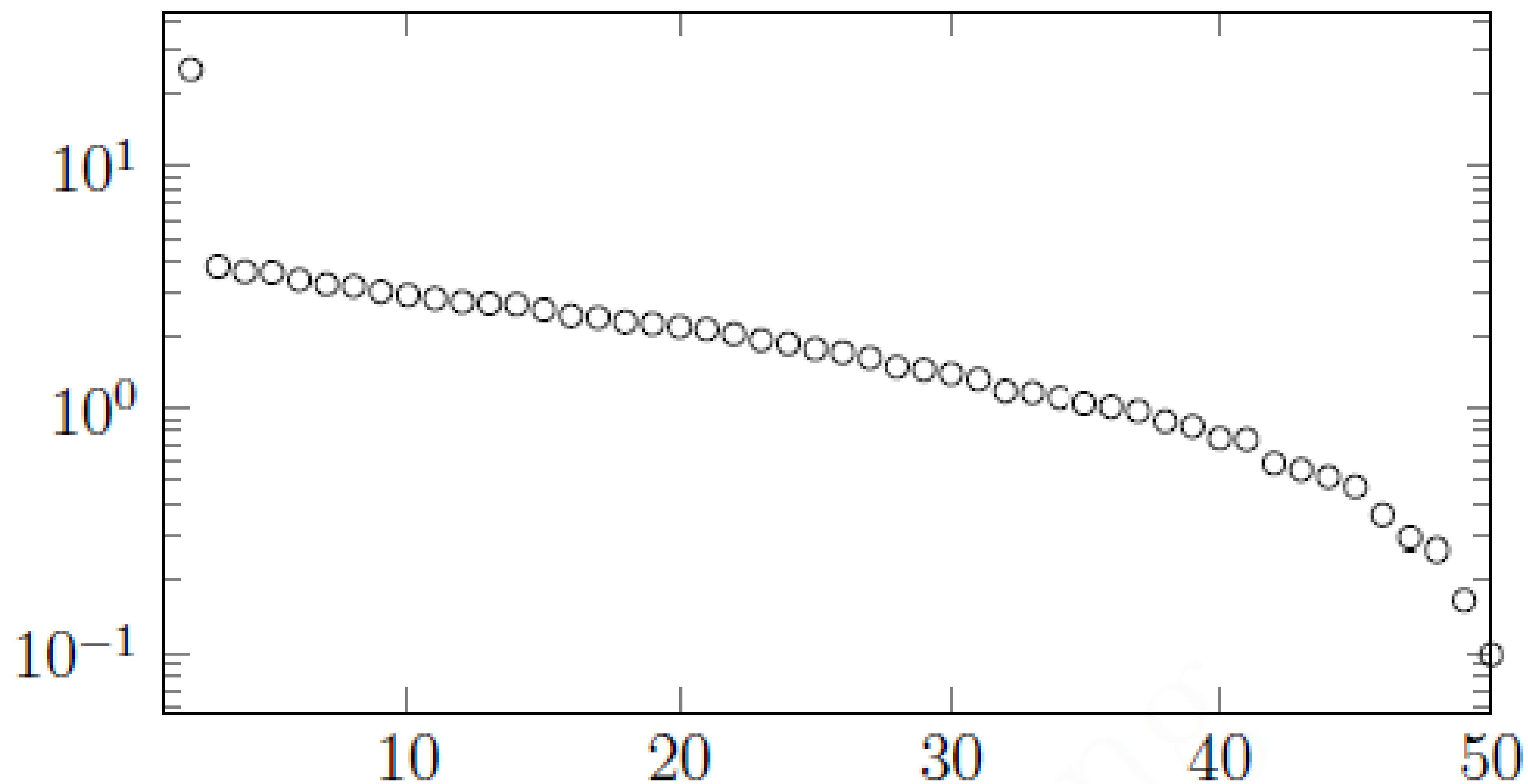
*Let $A = U\Sigma V^\top$ be an $m \times n$ matrix and let $A_k = U\Sigma_k V^\top$ where $\Sigma_k$ is same as $\Sigma$ except that the $(j, j)$ entry is $0$ for $j > k$. Then*

$$\min_{\text{rank}(S) \le k} \|A - S\|_F = \|A - A_k\|_F = (\sigma_{k+1}^2 + \cdots + \sigma_n^2)^{1/2}.$$

*In other words, $A_k$ is the best rank $k$ approximation of $A$ in Frobenius norm.*

# Example

- Generate a random $50 \times 50$ matrix $A$ using Julia
- Check the rank of matrix $A$, which should be $64$ most of the case.
- Plot the singular values:



# Applications

- Eigenvalues with PCA;
- Eigenvalues for extracting information from graph;
- SVD with recommender systems.