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

- 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

- 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

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

Contents

Definition

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

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

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

Implications

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

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

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

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

- The following statements are equivalent:
  - λ 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 λ.
- All nonzero vectors in $\mathcal{N}(A - \lambda I)$ are eigenvectors corresponding to λ.

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

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

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

Characteristic polynomial

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

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 λ_1 are nonzero solutions of

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

The eigenvectors belonging to λ_2 are nonzero solutions of

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

Example

Find the eigenvalues and the corresponding eigenvectors of

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

Example

Find the eigenvalues and the corresponding eigenvectors of

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

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 λ_i is the root of $p(\lambda)$.

- For real valued matrices,
  - Complex eigenvalues occur in **conjugate pairs**, i.e., if λ is an eigenvalue, so is $\bar{\lambda}$.
  - If $\mathbf{z}$ is an eigenvector belonging to a complex eigenvalue λ, 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 λ_i is the root of $p(\lambda)$.

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

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 λ_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 λ_k has *multiplicity* m_k. We know that $\sum_k m_k = n$.
- The characteristic polynomial can be written as

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

- 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 λ_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 λ_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

- 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:

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, λ is an eigenvalue of A if and only if λ^{-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.
- If A is invertible, λ is an eigenvalue of A if and only if λ^{-1} is an eigenvalue of A^{-1}.
- Proof:

Stochastic Matrix

- 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 ($\mathbf{1}^\top A = \mathbf{1}^\top$).
- 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 λ of A have $|\lambda| \leq 1$.
- Proof:

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 Λ is a real diagonal matrix.
- - The diagonal entries λ_i of Λ are eigenvalues of A.
  - The columns $\mathbf{q}_i$ of Q are eigenvectors belonging to λ_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 Λ is a real diagonal matrix.
- - The diagonal entries λ_i of Λ are eigenvalues of A.
  - The columns $\mathbf{q}_i$ of Q are eigenvectors belonging to λ_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 Λ is a real diagonal matrix.
-
  - The diagonal entries λ_i of Λ are eigenvalues of A.
  - The columns $\mathbf{q}_i$ of Q are eigenvectors belonging to λ_i, respectively.

Spectral Theorem

We can also write

$$
\begin{aligned}
A &= \begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^\top \\ \vdots \\ \mathbf{q}_n^\top \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 \mathbf{q}_1^\top \\ \vdots \\ \lambda_n \mathbf{q}_n^\top \end{bmatrix} \\
&= \sum_{i=1}^{n} \lambda_i \mathbf{q}_i \mathbf{q}_i^\top
\end{aligned}
$$

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;
- Σ is an $m \times n$ matrix whose off-diagonal entries are all 0's, and whose diagonal entries σ_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 Σ with respect to the above bases of $\mathbb{R}^m$ and $\mathbb{R}^n$.
- In other words,

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

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

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

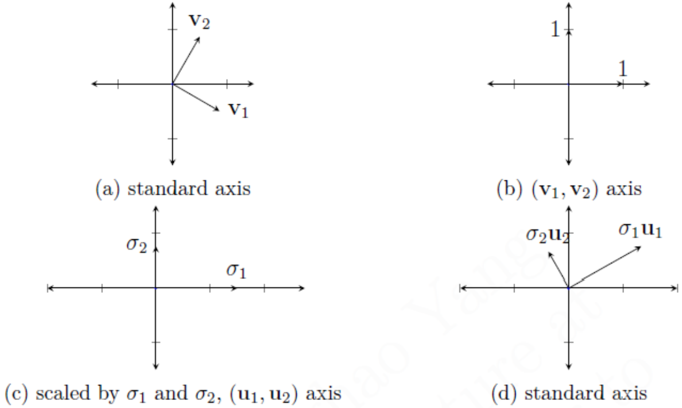
$$[\mathbf{x}]_V = V^\top \mathbf{x}$$

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 Σ with respect to the above bases of $\mathbb{R}^m$ and $\mathbb{R}^n$.
- In other words,

$$
\begin{aligned}
L(\mathbf{x}) &= U\Sigma V^\top \mathbf{x} \\
[L(\mathbf{x})]_U &= \Sigma [\mathbf{x}]_V \\
L(\mathbf{x}) &= U \\
[\mathbf{x}]_V &= V^\top \mathbf{x}
\end{aligned}
$$

Visualization of SVD



(a) standard axis

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

(c) scaled by σ_1 and σ_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 $\mathbf{x}_i$

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

  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

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

Theorem

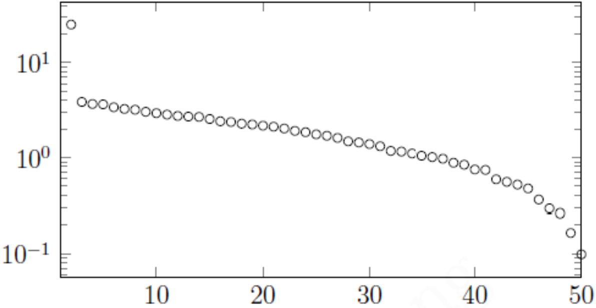
Let $A = U\Sigma V^\top$ be an $m \times n$ matrix and let $A_k = U\Sigma_k V^\top$ where Σ_k is same as Σ except that the (j,j) entry is 0 for $j > k$. Then

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

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

Example

- Generate a random 50×50 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.