Principles of Machine Learning

Lecture 2: Matrix Decomposition

Sharif University of Technology
Dept. of Aerospace Engineering

February 25, 2025

Table of Contents

Outline

Applications in ML & Data Analysis

- Dimensionality Reduction (e.g., PCA)
- Optimization (Linear Systems)
- Probabilistic Models (Normalization)
- Data Compression (Low-Rank Approximations)

Outline

Determinant: Definition

Definition (Determinant)

For a square matrix A, the determinant $\det(A)$ is a scalar that summarizes key properties of A:

- A is invertible if and only if $\det(A) \neq 0$.
- $|\det(A)|$ gives the scaling factor of volume under the transformation.
- The sign of $\det(A)$ indicates orientation (preserved or reversed).

Geometric Interpretation

Example (Volume Scaling)

For a 3×3 matrix A with column vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$,

$$|\det(A)| \quad \text{equals the volume of the parallelepiped spanned by } \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3.$$

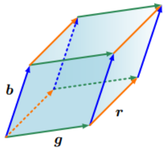

Figure: Parallelepiped volume

Laplace Expansion

Definition (Laplace Expansion)

The determinant of A can be computed by expanding along row i:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}),$$

where A_{ij} is the submatrix formed by deleting row i and column j.

$ad - bc$ $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Example

Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}.$$

Form the matrix

$$A = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 0 & -1 \\ 3 & 1 & 5 \end{bmatrix}.$$

Using the determinant formula, one finds $|\det(A)| = 47$.

Example 2: Laplace Expansion

Example

Let

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 3 & 1 & 0 \\ 1 & 4 & 5 \end{bmatrix}.$$

Expanding along the first row:

$$\det(A) = 2 \cdot \det \begin{bmatrix} 1 & 0 \\ 4 & 5 \end{bmatrix} - 0 + 1 \cdot \det \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix}.$$

With $\det \begin{bmatrix} 1 & 0 \\ 4 & 5 \end{bmatrix} = 5$ and $\det \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix} = 11$, we obtain:

$$\det(A) = 2 \cdot 5 + 11 = 21.$$

Properties: Determinant

- Multiplicative: $\det(AB) = \det(A)\det(B)$
- Transpose: $\det(A^\top) = \det(A)$
- Row swaps: Multiply by -1
- Row scaling: Multiply by the scalar factor
- addition of a Row with another Row: No change
- Invertibility: A invertible $\Leftrightarrow \det(A) \neq 0$

Properties: Additional Determinant Facts

- Block Diagonal: For $A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$, $\det(A) = \det(B)\det(C)$
- Scaling: $\det(kA) = k^n \det(A)$ for an $n \times n$ matrix

$$\begin{bmatrix} 3\times3 & 0 & 0 \\ 0 & 2\times2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Trace: Definition

Definition (Trace)

For a square matrix $A = [a_{ij}]$, the trace is defined as:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Trace: Properties

- Linearity: $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$
- Cyclic: $\text{tr}(AB) = \text{tr}(BA)$
- Similarity invariance: $\text{tr}(P^{-1}AP) = \text{tr}(A)$

Concept Map

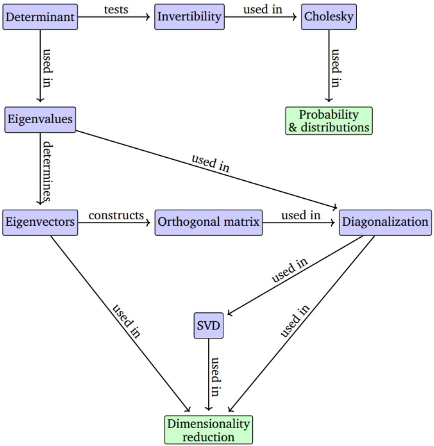

Figure: Mind Map of Concepts

Outline

Eigenvalue Definition

Definition (Eigenvalue & Eigenvector)

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

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

λx

x

The vector $\mathbf{x}$ is an **eigenvector** for λ.

Geometric Interpretation

Eigenvectors correspond to directions invariant under the linear transformation A. Eigenvalues scale eigenvectors in those directions.

Characteristic Polynomial

Theorem (Characteristic Equation)

The eigenvalues of A are roots of the ***characteristic polynomial****:*

$$p(\lambda) = \det(A - \lambda I)$$

Degree & Coefficients

For an $n \times n$ matrix:

- Degree n polynomial
- Leading term $(-1)^n \lambda^n$
- Trace: $\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i$
- Determinant: $\det(A) = \prod_{i=1}^{n} \lambda_i$

$Ax - \lambda x = 0$

$(A - \lambda I)\, x = 0$

$\underbrace{\text{rank}}_{< n} + \underbrace{\text{Nullity}}_{> 0} =$ تعداد ستون‌ها

Eigenvalue Computation

1. Solve characteristic equation $\det(A - \lambda I) = 0$
2. For each eigenvalue λ_i:
   - Solve $(A - \lambda_i I)\mathbf{x} = 0$
   - Solution space: Eigenspace E_{λ_i}

Example

For $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$:

$$\det(A - \lambda I) = (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0$$
$$\Rightarrow \lambda = 1, 3$$

$$\text{For } \lambda = 1: \quad (A - I)\mathbf{x} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\mathbf{x} = 0 \Rightarrow \mathbf{x} = t\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\text{For } \lambda = 3: \quad (A - 3I)\mathbf{x} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\mathbf{x} = 0 \Rightarrow \mathbf{x} = t\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Algebraic vs Geometric Multiplicity

Definition

- **Algebraic multiplicity**: Multiplicity of λ as a root of $p(\lambda)$
- **Geometric multiplicity**: $\dim E_\lambda = \dim \ker(A - \lambda I)$

Theorem

For any eigenvalue λ:

$$1 \leq \textit{Geometric multiplicity} \leq \textit{Algebraic multiplicity}$$

Diagonalizability

An $n \times n$ matrix is diagonalizable iff the geometric multiplicity equals algebraic multiplicity for all eigenvalues.

Applications

- **Principal Component Analysis**: Covariance matrix eigenvectors
- **Quantum Mechanics**: Observables as Hermitian operators
- **Vibration Analysis**: Natural frequencies in mechanical systems
- **PageRank Algorithm**: Dominant eigenvector of web matrix

Stability Analysis

In dynamical systems $\mathbf{x}' = A\mathbf{x}$:

- Stable if all $\text{Re}(\lambda_i) < 0$
- Unstable if any $\text{Re}(\lambda_i) > 0$

Example 1: 2×2 Matrix

Example

- $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$
- Characteristic eq:

$$\det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = (2-\lambda)^2 - 1 = 0$$

- $\Rightarrow 2-\lambda = \pm 1 \Rightarrow \lambda_1 = 1,\ \lambda_2 = 3$

Example: Eigenvectors

Example

- For $\lambda_1 = 1$:

$$(A - I) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \Rightarrow \quad x_1 + x_2 = 0.$$

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

- For $\lambda_2 = 3$:

$$(A - 3I) = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \quad \Rightarrow \quad -x_1 + x_2 = 0.$$

Choose $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Transformation Matrices

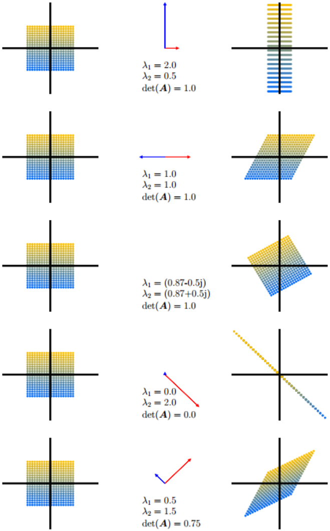

Figure: Eigenvalue-based transformation matrices

Eigenvectors

Eigenvectors point in the same direction (or opposite) after the transformation

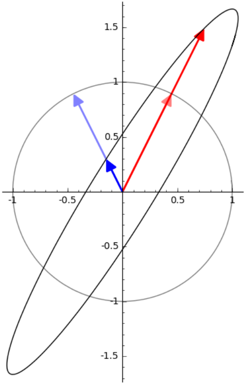

Figure: Eigenvectors do not change direction under a transformation

Determinant and Trace via Eigenvalues

Theorem (Determinant and Trace via Eigenvalues)

- *If A has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then:*

$$\det(A) = \prod_{i=1}^{n} \lambda_i \quad \textit{and} \quad \operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i.$$

- *These relationships simplify the calculation of these invariants when the eigenvalues are known.*

Definite Matrices

Definition (Positive Definite)

A symmetric matrix A is *positive definite* if

$$\mathbf{x}^\top A\mathbf{x} > 0, \quad \forall \mathbf{x} \neq 0.$$

- All eigenvalues > 0
- A is invertible; unique Cholesky decomposition exists

Definition (Negative Definite)

A is *negative definite* if

$$\mathbf{x}^\top A\mathbf{x} < 0, \quad \forall \mathbf{x} \neq 0.$$

- All eigenvalues < 0
- A is invertible; useful in concavity and stability analysis

Outline

Why Cholesky Decomposition?

- Many problems in numerical linear algebra involve solving systems of equations of the form $A\mathbf{x} = \mathbf{b}$, where A is a symmetric positive definite (SPD) matrix.
- Cholesky Decomposition is especially efficient for SPD matrices.
- It allows us to decompose A into a product of a lower triangular matrix and its transpose, reducing computational complexity compared to a full LU decomposition.
- This decomposition also improves numerical stability.

Cholesky: Def.

Definition (Cholesky Factorization)

For a symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower-triangular matrix L with positive diagonal entries such that

$$A = L\,L^\top.$$

L is the **Cholesky factor**.

Cholesky

Consider a 3×3 SPD matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

$$L = \begin{bmatrix} \ell_{11} & 0 & 0 \\ \ell_{21} & \ell_{22} & 0 \\ \ell_{31} & \ell_{32} & \ell_{33} \end{bmatrix}$$

such that $A = L\,L^\top$. The relationships are:

- $a_{11} = \ell_{11}^2 \quad \Rightarrow \quad \ell_{11} = \sqrt{a_{11}}$.
- $a_{12} = \ell_{11}\,\ell_{21} \quad \Rightarrow \quad \ell_{21} = \dfrac{a_{12}}{\ell_{11}}$.
- $a_{13} = \ell_{11}\,\ell_{31} \quad \Rightarrow \quad \ell_{31} = \dfrac{a_{13}}{\ell_{11}}$.
- $a_{22} = \ell_{21}^2 + \ell_{22}^2 \quad \Rightarrow \quad \ell_{22} = \sqrt{a_{22} - \ell_{21}^2}$.
- $a_{23} = \ell_{21}\,\ell_{31} + \ell_{22}\,\ell_{32} \quad \Rightarrow \quad \ell_{32} = \dfrac{a_{23} - \ell_{21}\,\ell_{31}}{\ell_{22}}$.

Cholesky: Determinant

- $A = L\,L^\top \quad \Rightarrow \quad \det(A) = (\det(L))^2$
- For lower-triangular L: $\det(L) = \ell_{11}\,\ell_{22}\,\cdots\,\ell_{nn}$
- Thus, $\det(A) = (\ell_{11}\,\ell_{22}\,\cdots\,\ell_{nn})^2$

Outline

Diagonal Matrices

Definition (Diagonal Matrix)

A diagonal matrix D has the form

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Since D is diagonal, $D^\top = D$. If $\lambda_i \neq 0$ for all i, then

$$D^{-1} = \begin{bmatrix} 1/\lambda_1 & 0 & \cdots & 0 \\ 0 & 1/\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\lambda_n \end{bmatrix}.$$

Diagonalization

Theorem (Spectral Decomposition)

If A has n linearly independent eigenvectors, then:

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

where D is diagonal matrix of eigenvalues and P has corresponding eigenvectors as columns.

Procedure

1. Find eigenvalues $\lambda_1, \ldots, \lambda_n$
2. Find corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$
3. Construct $P = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ and $D = \mathrm{diag}(\lambda_i)$

Theorem (Spectral Theorem)

For real symmetric matrix A:

- *All eigenvalues are real*
- *Eigenvectors can be chosen orthogonal*
- $A = QDQ^\top$ *where Q is orthogonal*

$Q^T AQ = D$

$Q^{-1} AQ = D$

$Q^{-1} = Q^T$

$Q = [v_1 \; v_2 \; \cdots \; v_n]$

Positive Definite (PD) Matrices

If all $\lambda_i > 0 \longrightarrow$ PD

$$\mathbf{x}^\top A\mathbf{x} > 0 \quad \forall \mathbf{x} \neq 0$$

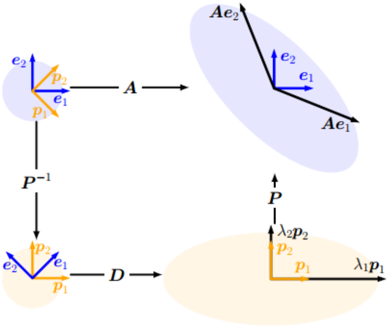

Figure: Sequential transformations: Rotation, scaling, reorientation.

Example (Diagonalize)

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$$

- Characteristic equation:

$$\det \begin{bmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{bmatrix} = (4-\lambda)(3-\lambda) - 2 = \lambda^2 - 7\lambda + 10 = 0.$$

- Eigenvalues: $\lambda_1 = 5$, $\lambda_2 = 2$.

Example

Find eigenvectors for

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}.$$

- For $\lambda_1 = 5$:

$$A - 5I = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix} \quad \Rightarrow \quad -x_1 + x_2 = 0 \quad \Rightarrow \quad x_2 = x_1.$$

Choose $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Example

- For $\lambda_2 = 2$:

$$A - 2I = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \quad \Rightarrow \quad 2x_1 + x_2 = 0 \quad \Rightarrow \quad x_2 = -2x_1.$$

Choose $v_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$.

Example

Form P and D.

- $P = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$.
- $D = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$.
- Thus, $A = P\,D\,P^{-1}$.

- $A = P\,D\,P^{-1}$.
- $\det(A) = \det(P)\,\det(D)\,\det(P^{-1})$.
- Since $\det(P^{-1}) = 1/\det(P)$, we obtain

$$\det(A) = \det(D).$$

Power: A^k

- $A = P D P^{-1}$.
- Thus, $A^k = (P D P^{-1})^k = P D^k P^{-1}$.
- $D^k = \text{diag}(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k)$.

$A^k = (PD\cancel{P^{-1}})(\cancel{P}D\cancel{P^{-1}})\cdots(\cancel{P}DP^{-1})$

$\boxed{A^k = P D^k P^{-1}}$

Outline

SVD: Motivation

- Decomposes any $m \times n$ matrix into three factors.
- Applications:
  - Dimensionality reduction (e.g., PCA)
  - Data compression, noise reduction
  - Pseudoinversion of ill-posed problems

SVD: Definition

Definition (Singular Value Decomposition)

For $A \in \mathbb{R}^{m \times n}$, the SVD is

$$A = U \Sigma V^\top,$$

where:

- U is an $m \times m$ orthogonal matrix (left singular vectors).
- Σ is an $m \times n$ rectangular diagonal matrix with nonnegative entries.
- V is an $n \times n$ orthogonal matrix (right singular vectors).

SVD: Dimensions

- For $A \in \mathbb{R}^{m \times n}$:
  - U: $m \times m$ (or $m \times r$ in reduced form)
  - Σ: $m \times n$ (or $r \times r$ in reduced form)
  - V: $n \times n$ (or $n \times r$ in reduced form)
- Σ contains singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$ with $p = \min(m, n)$.

$$\Sigma = \left[\begin{array}{cccc|c} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & 0 \\ & & & \sigma_p & \end{array}\right]$$

SVD: Roles

- $V^\top$: Rotates input into a new coordinate system.
- Σ: Scales the coordinates (by singular values).
- U: Rotates the scaled data into the output space.

Σ: Structure

- Σ is $m \times n$ and "diagonal" (nonzero only on the main diagonal).
- Its form:

$$\Sigma = \left[\begin{array}{cccc} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p \\ \hline 0 & 0 & \cdots & 0 \end{array}\right], \quad p = \min(m, n).$$

SVD: Fundamental Theorem

Theorem (Singular Value Decomposition)

For any real $m \times n$ matrix A, there exist orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$, and a diagonal matrix $\Sigma \in \mathbb{R}^{m \times n}$ with non-negative entries, such that:

$$A = U \Sigma V^\top$$

*The diagonal entries $\sigma_i = \Sigma_{ii}$ are called **singular values**.*

Key Properties

- $\text{rank}(A)$ = number of non-zero singular values
- $\|A\|_F = \sqrt{\sum \sigma_i^2}$ (Frobenius norm)
- $\|A\|_2 = \sigma_1$ (Spectral norm)

Singular Value Computation

1. Compute $A^\top A$ (symmetric $n \times n$ matrix)
2. Solve the eigenvalue problem:

$$A^\top A v_i = \lambda_i v_i$$

3. Extract singular values:

$$\sigma_i = \sqrt{\lambda_i} \quad \text{with} \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

$A = U\Sigma V^T$

$A^T = V\Sigma^T U^T$

$= V\Sigma^T \cancel{U^T U} \Sigma V^T$

$A^T A = (V)(\Sigma^T\Sigma)(V^T)$

P

Theorem (Spectral Guarantee)

For any real matrix A:

- *$A^\top A$ is positive semi-definite*
- *All $\lambda_i \geq 0$*
- *$\exists$ orthonormal eigenbasis $\{v_i\}$ for $A^\top A$*

SVD Construction Procedure

1. Compute eigendecomposition:

$$A^\top A = V\Lambda V^\top, \quad \Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$$

2. Form singular value matrix:

$$\Sigma = \begin{pmatrix} \mathrm{diag}(\sigma_1, \dots, \sigma_r) & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{m \times n}, \quad \sigma_i = \sqrt{\lambda_i}$$

3. Construct orthogonal U:

$$u_i = \frac{1}{\sigma_i} A v_i \quad (i = 1, \dots, r)$$

Complete to orthonormal basis for $\mathbb{R}^m$

Example

For $A = \begin{pmatrix} 3 & 0 \\ 0 & -4 \end{pmatrix}$:

$$A^\top A = \begin{pmatrix} 9 & 0 \\ 0 & 16 \end{pmatrix}$$

$$\sigma_1 = 4, \ \sigma_2 = 3$$

$$V = I, \ U = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

SVD: Intuition

- Interpret A as a sequence:
  1. $V^\top$: Reorients the input.
  2. Σ: Scales the reoriented data.
  3. U: Maps the result into the output space.

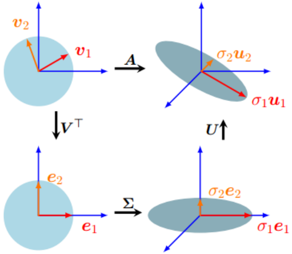

Figure: Sequential transformations in SVD.

Example

Consider

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \in \mathbb{R}^{3\times 2}.$$

- **Step 1:** Compute $A^\top A$:

$$A^\top A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix}.$$

- The eigenvalues of $A^\top A$ (say, λ_1, λ_2) yield singular values: $\sigma_i = \sqrt{\lambda_i}$.

Example

- Assume the computed singular values are:

$$\sigma_1 \approx 9.5255, \quad \sigma_2 \approx 0.5143.$$

- **Step 2:** Form V using the eigenvectors of $A^\top A$. For example,

$$V \approx \begin{bmatrix} -0.6196 & -0.7849 \\ -0.7849 & 0.6196 \end{bmatrix}.$$

- **Step 3:** Compute U as

$$U = A\,V\Sigma^{-1}.$$

- In the reduced SVD, U is 3×2 and Σ is 2×2.

Low-Rank Approximation Theory

Theorem (Eckart–Young (1936))

For any matrix A with SVD $A = \sum_{i=1}^{r} \sigma_i u_i v_i^\top$, the best rank-$k$ approximation is:

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^\top$$

This minimizes both:

$$\|A - B\|_F^2 = \sum_{i=k+1}^{r} \sigma_i^2 \quad \text{and} \quad \|A - B\|_2 = \sigma_{k+1}$$

over all rank-k matrices B.

Applications

Applications

- Image compression (JPEG)
- Recommendation systems
- Dimensionality reduction (PCA)
- Noise reduction in signal processing

SVD vs EVD: Deep Comparison

Eigenvalue Decomposition (EVD)	**Singular Value Decomposition (SVD)**
Requires square matrix	Works for any rectangular matrix
$A = PDP^{-1}$	$A = U\Sigma V^\top$
May contain complex values	Always real, non-negative σ_i
Sensitive to non-normality	Numerically stable
Requires full rank for diagonalization	Always exists
Reveals operator geometry	Reveals input-output geometry

Theorem (SVD-EVD Connection)

For normal matrices ($AA^\top = A^\top A$), SVD coincides with EVD:

$$\sigma_i = |\lambda_i|, \;\; U = V \; (\textit{up to sign})$$

Advanced Applications

Matrix Pseudoinverse

$$A^+ = V\Sigma^+ U^\top, \quad \Sigma^+_{ii} = \begin{cases} 1/\sigma_i & \sigma_i \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Minimum norm least squares solution: $\mathbf{x}^* = A^+\mathbf{b}$

Numerical Stability

Condition number $\kappa(A) = \sigma_{\max}/\sigma_{\min}$ determines:

- Matrix invertibility
- Stability of linear systems Sensitivity to numerical errors