Singular Value Decomposition

Sergio Peignier

Definition of the SVD

Singular Value Decomposition

For any matrix $A \in \mathbb{R}^{m \times n}$, there exist matrices

$$U \in \mathbb{R}^{m \times m}, \quad \Sigma \in \mathbb{R}^{m \times n}, \quad V \in \mathbb{R}^{n \times n}$$

such that

$$\boxed{A = U \Sigma V^\top}$$

- The SVD exists for *any* matrix (square or rectangular).
- Unlike eigenvalue decomposition no diagonalisability is required.
- Singular values are uniquely determined (up to ordering)
- Singular vectors are unique (up to phase/sign) when singular values are distinct

Left and Right Singular Vectors

- U: unitary matrix, columns are **left singular vectors**
- V: unitary matrix, columns are **right singular vectors**

Unitary Matrix

A matrix Q is unitary if

$$Q^\top Q = QQ^\top = I$$

- Columns are orthonormal
- Lengths and inner products are preserved:

$$\|Qx\|_2 = \|x\|_2, \qquad \langle Qx, Qy \rangle = \langle x, y \rangle$$

Singular Values

- Σ is a rectangular diagonal matrix
- Nonzero entries lie on the main diagonal

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0, \qquad p = \min(m, n)$$

Interpretation

Each singular value σ_i measures how much A stretches vectors along the corresponding singular direction.

Alternative Characterization

A scalar $\sigma \geq 0$ is a singular value of A if and only if there exist unit vectors $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$ such that

$$Av = \sigma u, \qquad A^\top u = \sigma v$$

- v: right singular vector
- u: left singular vector

Positive Semi-Definite Matrix

To understand why singular values are real and non-negative, consider:

$$A^T A \quad (n \times n)$$

- $A^T A$ is symmetric
- For any vector $x \in \mathbb{R}^n$:

$$x^T (A^T A) x = (Ax)^T (Ax) = \|Ax\|^2 \geq 0$$

- Any matrix M with $x^T M x \geq 0 \; \forall x$ is called **positive semi-definite (PSD)**

Takeaway: PSD matrices always have **non-negative eigenvalues**, which is key for singular values.

Diagonalize the PSD Matrix

- Since $A^T A$ is symmetric and PSD, it can be diagonalized:

$$A^T A = V \Lambda V^T$$

  where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \geq 0$
- Eigenvalues λ_i are real and non-negative
- Columns of V are orthonormal eigenvectors

Observation: This is the first step toward SVD: - Eigenvectors of $A^T A \rightarrow$ right singular vectors - Square roots of eigenvalues $\rightarrow$ singular values

Construct the SVD

- Define singular values:

$$\sigma_i = \sqrt{\lambda_i} \geq 0$$

- Form the diagonal matrix Σ with σ_i on the diagonal
- Define left singular vectors:

$$U = AV\Sigma^{-1} \quad \text{(columns of } U \text{ are orthonormal)}$$

- Then

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

Construct the SVD

- Define singular values:

$$\sigma_i = \sqrt{\lambda_i} \geq 0$$

- Form the diagonal matrix Σ with σ_i on the diagonal
- Define left singular vectors:

$$U = AV\Sigma^{-1} \quad \text{(columns of } U \text{ are orthonormal)}$$

- Then

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

Key Points:

- SVD exists for **any matrix**
- Singular values are **real and non-negative**
- Provides a clean geometric decomposition: rotation $\rightarrow$ scaling $\rightarrow$ rotation

Relation to Eigenvalue Decomposition

$$A^\top A = V(\Sigma^\top \Sigma)V^\top, \qquad AA^\top = U(\Sigma\Sigma^\top)U^\top$$

- Columns of V: eigenvectors of $A^\top A$
- Columns of U: eigenvectors of $AA^\top$
- $\sigma_i = \sqrt{\lambda_i(A^\top A)}$

Complex-Valued Matrices

- The SVD exists for any $A \in \mathbb{C}^{m \times n}$
- Transposes are replaced by conjugate transposes

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

- U, V are unitary
- Singular values remain real and non-negative

Geometric Interpretation

Consider the unit sphere

$$S = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$$

- Applying A maps the sphere to a hyper-ellipse
- SVD explains this transformation in three steps

Three-Step Decomposition

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

1. $V^\top$: rotation of the input space
2. Σ: stretching along orthogonal axes
3. U: rotation in the output space

sphere $\rightarrow$ stretched ellipse $\rightarrow$ rotated ellipse

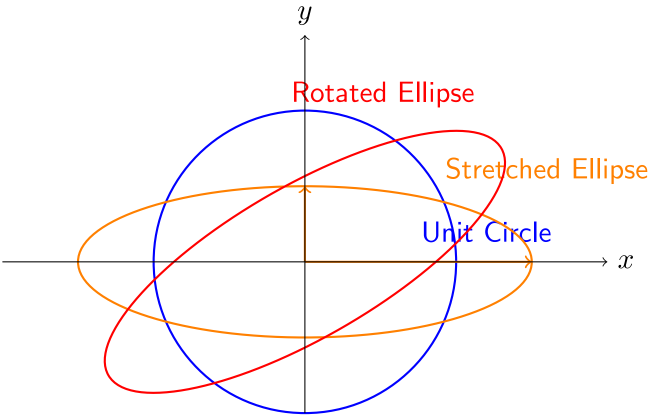
What Does the SVD Tell Us?

In which directions does a matrix stretch space, and by how much?

rotation $\rightarrow$ scaling along orthogonal directions $\rightarrow$ rotation

- Directions of maximal amplification: v_i
- Amount of amplification: σ_i
- Output directions: u_i

Geometric Interpretation of SVD (2D)



- V: rotate to circle to singular directions
- Σ: stretches along singular directions
- U: rotates ellipse to final orientation

Reduced SVD

Assume $\text{rank}(A) = r \leq \min(m, n)$.

$$A = \hat{U}\hat{\Sigma}\hat{V}^\top$$

- $\hat{U} \in \mathbb{R}^{m \times r}$
- $\hat{V} \in \mathbb{R}^{n \times r}$
- $\hat{\Sigma} \in \mathbb{R}^{r \times r}$

Practical Importance

The reduced SVD is cheaper to compute and commonly used in applications.

Full SVD

To obtain full unitary matrices:

- Extend $\hat{U}$ and $\hat{V}$ to full orthonormal bases
- Append zeros to $\hat{\Sigma}$

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

- $U \in \mathbb{R}^{m\times m}$
- $V \in \mathbb{R}^{n\times n}$

Summary

- SVD exists for any matrix
- Singular values are real and non-negative
- SVD separates geometry into rotations and scalings
- Provides orthonormal bases aligned with the action of A

$$\boxed{A = U\Sigma V^\top}$$