Lecture 04
Singular Value Decomposition (SVD)

Songting Luo

Department of Mathematics
Iowa State University

MATH 562 Numerical Analysis II

Outline

1. Singular Value Decomposition (SVD)

Outline

Geometric Observation

- The image of unit sphere under any $m \times n$ matrix is a hyperellipse
- Given a unit sphere $S \in \mathbb{R}^n$, let $\mathbf{A}S$ denote the shape after transformation
- SVD is

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$$

  where $\mathbf{U} \in \mathbb{C}^{m\times m}$ and $\mathbf{V} \in \mathbb{C}^{n\times n}$ and $\mathbf{\Sigma} \in \mathbf{R}^{m\times n}$ is diagonal.
- Singular values are diagonal entries of $\mathbf{\Sigma}$, correspond to the principal semiaxes, with entries $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n \geqslant 0$.
- Left singular vectors of $\mathbf{A}$ are column vectors of $\mathbf{U}$ and are oriented in the directions of the principal semiaxes of $\mathbf{A}S$
- Right singular vectors of $\mathbf{A}$ are column vectors of $\mathbf{V}$ and are the preimages of the principal semiaxes of $\mathbf{A}S$
- $\mathbf{A}\mathbf{v}_j = \sigma_j \mathbf{u}_j$ for $1 \leqslant j \leqslant n$

Two Different Types of SVD

- Full SVD: $\mathbf{U} \in \mathbb{C}^{m\times m}$, $\mathbf{\Sigma} \in \mathbb{R}^{m\times n}$ and $\mathbf{V} \in \mathbb{C}^{n\times n}$ is

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$$

Two Different Types of SVD

- Full SVD: $\mathbf{U} \in \mathbb{C}^{m\times m}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{m\times n}$ and $\mathbf{V} \in \mathbb{C}^{n\times n}$ is

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$$

- Reduced SVD: $\hat{\mathbf{U}} \in \mathbb{C}^{m\times n}$, $\hat{\boldsymbol{\Sigma}} \in \mathbb{R}^{n\times n}$ and $\mathbf{V} \in \mathbb{C}^{n\times n}$ (assume $m \geqslant n$) is

$$\mathbf{A} = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}\mathbf{V}^*$$

Two Different Types of SVD

- Full SVD: $\mathbf{U} \in \mathbb{C}^{m\times m}$, $\mathbf{\Sigma} \in \mathbb{R}^{m\times n}$ and $\mathbf{V} \in \mathbb{C}^{n\times n}$ is

$$\mathbf{A} = \mathbf{U\Sigma V}^*$$

- Reduced SVD: $\hat{\mathbf{U}} \in \mathbb{C}^{m\times n}$, $\hat{\mathbf{\Sigma}} \in \mathbb{R}^{n\times n}$ and $\mathbf{V} \in \mathbb{C}^{n\times n}$ (assume $m \geqslant n$) is

$$\mathbf{A} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\mathbf{V}^*$$

- What if $m \leqslant n$?
- Furthermore, notice that

$$\mathbf{A} = \sum_{i=1}^{\min\{m,n\}} \sigma_i \mathbf{u}_i \mathbf{v}_i^*$$

so we can keep only entries of $\mathbf{U}$ and $\mathbf{V}$ corresponding to nonzero σ_i.

Existence of SVD

Theorem (Existence)

Every matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ *has an SVD.*

Proof

Let $\sigma_1 = \|\mathbf{A}\|_2$. There exists $\mathbf{v}_1 \in \mathbb{C}^n$ with $\|\mathbf{v}\|_1 = 1$ and $\|\mathbf{A}\mathbf{v}_1\|_2 = \sigma_1$. Let $\mathbf{U}_1$ and $\mathbf{V}_1$ be unitary matrices whose first columns are $\mathbf{u}_1 = \mathbf{A}\mathbf{v}_1/\sigma_1$ (or any unit-length vector if $\sigma_1 = 0$) and $\mathbf{v}_1$, respectively. Note that,

$$\mathbf{U}_1^*\mathbf{A}\mathbf{V}_1 = \mathbf{S} = \begin{bmatrix} \sigma_1 & \boldsymbol{\omega}^* \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$$

Futhermore, $\boldsymbol{\omega} = \mathbf{0}$ because $\|\mathbf{S}\|_2 = \sigma_1$, and

$$\left\| \begin{bmatrix} \sigma_1 & \boldsymbol{\omega}^* \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \boldsymbol{\omega} \end{bmatrix} \right\|_2 \geqslant \sigma_1^2 + \boldsymbol{\omega}^*\boldsymbol{\omega} = \sqrt{\sigma_1^2 + \boldsymbol{\omega}^*\boldsymbol{\omega}} \left\| \begin{bmatrix} \sigma_1 \\ \boldsymbol{\omega} \end{bmatrix} \right\|_2$$

implying that $\sigma_1 \geqslant \sqrt{\sigma_1^2 + \boldsymbol{\omega}^*\boldsymbol{\omega}}$ and $\boldsymbol{\omega} = \mathbf{0}$.

Existence of SVD Cont'd

We then prove by induction. If $m = 1$ or $n = 1$, then $\mathbf{B}$ is empty and we have $\mathbf{A} = \mathbf{U}_1\mathbf{S}\mathbf{V}_1^*$. Otherwise, suppose $\mathbf{B} = \mathbf{U}_2\mathbf{\Sigma}_2\mathbf{V}_2^*$, and then

$$\mathbf{A} = \underbrace{\mathbf{U}_1 \begin{bmatrix} 1 & \mathbf{0}^* \\ \mathbf{0} & \mathbf{U}_2 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} \sigma_1 & \mathbf{0}^* \\ \mathbf{0} & \mathbf{\Sigma}_2 \end{bmatrix}}_{\mathbf{\Sigma}} \underbrace{\begin{bmatrix} 1 & \mathbf{0}^* \\ \mathbf{0} & \mathbf{V}_2^* \end{bmatrix} \mathbf{V}_1^*}_{\mathbf{V}^*}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary.

Uniqueness of SVD

Theorem (Uniqueness)

The singular values $\{\sigma_j\}$ are uniquely determined. If $\mathbf{A}$ *is square and the* σ_j *are distinct, the left and right singular vectors are uniquely determined up to complex signs (i.e., complex scalar factors of absolute value* 1*).*

Uniqueness of SVD

Theorem (Uniqueness)

The singular values $\{\sigma_j\}$ are uniquely determined. If $\mathbf{A}$ *is square and the σ_j are distinct, the left and right singular vectors are uniquely determined up to complex signs (i.e., complex scalar factors of absolute value 1).*

Geometric argument: If the lengths of semiaxes of a hyperellipse are distinct, then the semiaxes themselves are determined by the geometry up to signs.

Uniqueness of SVD Cont?d

Algebraic argument: Based on 2-norm and prove by induction. Consider the case where the σ_j are distinct. The 2-norm is unique, so is σ_1. If $\mathbf{v}_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $\mathbf{A}$, implying that σ_1 is not a simple singular value.

Once $\sigma_1, \mathbf{u}_1$, and $\mathbf{v}_1$ are determined, the remainder of SVD is determined by the space orthogonal to $\mathbf{v}_1$. Because $\mathbf{v}_1$ is unique up to sign, the orthogonal subspace is uniquely defined. Then prove by induction.

Uniqueness of SVD Cont?d

Algebraic argument: Based on 2-norm and prove by induction. Consider the case where the σ_j are distinct. The 2-norm is unique, so is σ_1. If $\mathbf{v}_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $\mathbf{A}$, implying that σ_1 is not a simple singular value.

Once $\sigma_1, \mathbf{u}_1$, and $\mathbf{v}_1$ are determined, the remainder of SVD is determined by the space orthogonal to $\mathbf{v}_1$. Because $\mathbf{v}_1$ is unique up to sign, the orthogonal subspace is uniquely defined. Then prove by induction.

- Question: What if we change the sign of a singular vector?

Uniqueness of SVD Cont?d

Algebraic argument: Based on 2-norm and prove by induction. Consider the case where the σ_j are distinct. The 2-norm is unique, so is σ_1. If $\mathbf{v}_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $\mathbf{A}$, implying that σ_1 is not a simple singular value.

Once $\sigma_1, \mathbf{u}_1$, and $\mathbf{v}_1$ are determined, the remainder of SVD is determined by the space orthogonal to $\mathbf{v}_1$. Because $\mathbf{v}_1$ is unique up to sign, the orthogonal subspace is uniquely defined. Then prove by induction.

- Question: What if we change the sign of a singular vector?
- Question: What if σ_i is not distinct?

SVD vs. Eigenvalue Decomposition

- Eigenvalue decomposition of nondefective matrix $\mathbf{A}$ is $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$

SVD vs. Eigenvalue Decomposition

- Eigenvalue decomposition of nondefective matrix $\mathbf{A}$ is $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$
- Differences between SVD and Eigenvalue Decomposition
  - Not every matrix has eigenvalue decomposition, but every matrix has singular value decomposition
  - Eigenvalues may not always be real numbers, but singular values are always non-negative real numbers
  - Eigenvectors are not always orthogonal to each other (orthogonal for symmetric matrices), but left (or right) singular vectors are orthogonal to each other

SVD vs. Eigenvalue Decomposition

- Eigenvalue decomposition of nondefective matrix $\mathbf{A}$ is $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$
- Differences between SVD and Eigenvalue Decomposition
  - Not every matrix has eigenvalue decomposition, but every matrix has singular value decomposition
  - Eigenvalues may not always be real numbers, but singular values are always non-negative real numbers
  - Eigenvectors are not always orthogonal to each other (orthogonal for symmetric matrices), but left (or right) singular vectors are orthogonal to each other
- Similarities
  - Singular values of $\mathbf{A}$ are square roots of eigenvalues of $\mathbf{A}\mathbf{A}^*$ and $\mathbf{A}^*\mathbf{A}$, and their eigenvectors are left and right singular vectors, respectively
  - Singular values of hermitian matrices are absolute values of eigenvalues, and eigenvectors are singular vectors (up to complex signs)
  - This relationship can be used to compute singular values by hand

Matrix Properties via SVD

- Let r be number of nonzero singular values of $\mathbf{A} \in \mathbb{C}^{m \times n}$
  - $rank(\mathbf{A}) = r$
  - $range(\mathbf{A}) = \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r \rangle$
  - $null(\mathbf{A}) = \langle \mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \ldots, \mathbf{v}_n \rangle$

Matrix Properties via SVD

- Let r be number of nonzero singular values of $\mathbf{A} \in \mathbb{C}^{m \times n}$
  - $rank(\mathbf{A}) = r$
  - $range(\mathbf{A}) = \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r \rangle$
  - $null(\mathbf{A}) = \langle \mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \ldots, \mathbf{v}_n \rangle$
- 2-norm and Frobenius norm
  - $\|\mathbf{A}\|_2 = \sigma_1$ and $\|\mathbf{A}\|_F = \sqrt{\sum_i \sigma_i^2}$

Matrix Properties via SVD

- Let r be number of nonzero singular values of $\mathbf{A} \in \mathbb{C}^{m \times n}$
  - $rank(\mathbf{A}) = r$
  - $range(\mathbf{A}) = \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r \rangle$
  - $null(\mathbf{A}) = \langle \mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \ldots, \mathbf{v}_n \rangle$
- 2-norm and Frobenius norm
  - $\|\mathbf{A}\|_2 = \sigma_1$ and $\|\mathbf{A}\|_F = \sqrt{\sum_i \sigma_i^2}$
- Determinant of matrix
  - For $\mathbf{A} \in \mathbb{C}^{m \times m}$, $|det(\mathbf{A})| = \prod_{i=1}^{m} \sigma_i$.

Matrix Properties via SVD

- Let r be number of nonzero singular values of $\mathbf{A} \in \mathbb{C}^{m \times n}$
  - $rank(\mathbf{A}) = r$
  - $range(\mathbf{A}) = \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r \rangle$
  - $null(\mathbf{A}) = \langle \mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \ldots, \mathbf{v}_n \rangle$
- 2-norm and Frobenius norm
  - $\|\mathbf{A}\|_2 = \sigma_1$ and $\|\mathbf{A}\|_F = \sqrt{\sum_i \sigma_i^2}$
- Determinant of matrix
  - For $\mathbf{A} \in \mathbb{C}^{m \times m}$, $|det(\mathbf{A})| = \prod_{i=1}^{m} \sigma_i$.
- However, SVD may not be the most efficient way in solving problems
- Algorithms for SVD are similar to those for eigenvalue decomposition and we will discuss them later in the semester