AMS526: Numerical Analysis I (Numerical Linear Algebra)

Lecture 4: Singular Value Decomposition

Xiangmin Jiao

SUNY Stony Brook

Geometric Observation

- The image of unit sphere under any $m \times n$ matrix is a *hyperellipse*
- Give a unit sphere S in $\mathbb{R}^n$, let $\boldsymbol{AS}$ denote the shape after transformation
- SVD is
$$\boldsymbol{A} = \boldsymbol{U\Sigma V}^*$$
where $\boldsymbol{U} \in \mathbb{C}^{m\times m}$ and $\boldsymbol{V} \in \mathbb{C}^{n\times n}$ is unitary and $\Sigma \in \mathbb{R}^{m\times n}$ is diagonal
- *Singular values* are diagonal entries of $\boldsymbol{\Sigma}$, correspond to the principal semiaxes, with entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.
- *Left singular vectors* of $\boldsymbol{A}$ are column vectors of $\boldsymbol{U}$ and are oriented in the directions of the principal semiaxes of $\boldsymbol{AS}$
- *Right singular vectors* of $\boldsymbol{A}$ are column vectors of $\boldsymbol{V}$ and are the preimages of the principal semiaxes of $\boldsymbol{AS}$
- $\boldsymbol{A}\boldsymbol{v}_j = \sigma_j \boldsymbol{u}_j$ for $1 \leq j \leq n$

Two Different Types of SVD

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

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

Two Different Types of SVD

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

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

- **Reduced SVD**: $\hat{\boldsymbol{U}} \in \mathbb{C}^{m\times n}$, $\hat{\boldsymbol{\Sigma}} \in \mathbb{R}^{n\times n}$ (assume $m \geq n$)

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

Two Different Types of SVD

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

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

- **Reduced SVD**: $\hat{\boldsymbol{U}} \in \mathbb{C}^{m\times n}$, $\hat{\boldsymbol{\Sigma}} \in \mathbb{R}^{n\times n}$ (assume $m \geq n$)

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

- Furthermore, notice that

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

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

Existence of SVD

Theorem

(Existence) Every matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ has an SVD.

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

$$\boldsymbol{U}_1^*\boldsymbol{A}\boldsymbol{V}_1 = \boldsymbol{S} = \begin{bmatrix} \sigma_1 & \boldsymbol{\omega}^* \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix}. \tag{1}$$

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

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

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

Existence of SVD Cont'd

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

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

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

Uniqueness of SVD

Theorem

(Uniqueness) The singular values $\{\sigma_j\}$ are uniquely determined. If $\boldsymbol{A}$ 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 $\boldsymbol{v}_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $\boldsymbol{A}$, implying that σ_1 is not a simple singular value.

Once $\sigma_1, \boldsymbol{u}_1$, and $\boldsymbol{v}_1$ are determined, the remainder of SVD is determined by the space orthogonal to $\boldsymbol{v}_1$. Because $\boldsymbol{v}_1$ 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 $\boldsymbol{v}_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $\boldsymbol{A}$, implying that σ_1 is not a simple singular value.

Once $\sigma_1, \boldsymbol{u}_1$, and $\boldsymbol{v}_1$ are determined, the remainder of SVD is determined by the space orthogonal to $\boldsymbol{v}_1$. Because $\boldsymbol{v}_1$ 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 $\boldsymbol{v}_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $\boldsymbol{A}$, implying that σ_1 is not a simple singular value.

Once $\sigma_1, \boldsymbol{u}_1$, and $\boldsymbol{v}_1$ are determined, the remainder of SVD is determined by the space orthogonal to $\boldsymbol{v}_1$. Because $\boldsymbol{v}_1$ 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 $\boldsymbol{A}$ is $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$

SVD vs. Eigenvalue Decomposition

- Eigenvalue decomposition of nondefective matrix $\boldsymbol{A}$ is $\boldsymbol{A} = \boldsymbol{X\Lambda X}^{-1}$
- 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 $\boldsymbol{A}$ is $\boldsymbol{A} = \boldsymbol{X\Lambda X}^{-1}$
- 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 $\boldsymbol{A}$ are square roots of eigenvalues of $\boldsymbol{AA}^*$ and $\boldsymbol{A}^*\boldsymbol{A}$, 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 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$
  - rank($\boldsymbol{A}$) is r
  - range($\boldsymbol{A}$) $= \langle \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r \rangle$
  - null($\boldsymbol{A}$) $= \langle \boldsymbol{v}_{r+1}, \boldsymbol{v}_{r+2}, \ldots, \boldsymbol{v}_n \rangle$
- 2-norm and Frobenius norm
  - $\|\boldsymbol{A}\|_2 = \sigma_1$ and $\|\boldsymbol{A}\|_F = \sqrt{\sum_i \sigma_i^2}$
- Determinant of matrix
  - For $\boldsymbol{A} \in \mathbb{C}^{m\times m}$, $|\det(\boldsymbol{A})| = \prod_{i=1}^m \sigma_i$

Matrix Properties via SVD

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