AMS526: Numerical Analysis I (Numerical Linear Algebra)

Lecture 5: Singular Value Decomposition

Xiangmin Jiao

SUNY Stony Brook

September 16, 2008

Review

- Matrix norms
  - 1-norm
  - ∞-norm
  - 2-norm
  - Frobenius norm

Invariance under Unitary Multiplication

Theorem

For any $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ *and unitary* $\boldsymbol{Q} \in \mathbb{C}^{m\times m}$, *we have*

$$\|\boldsymbol{QA}\|_2 = \|\boldsymbol{A}\|_2 \text{ and } \|\boldsymbol{QA}\|_F = \|\boldsymbol{A}\|_F.$$

In other words, 2-norm and Frobenius norms are invariant under unitary multiplication.

Invariance under Unitary Multiplication

Theorem

For any $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ *and unitary* $\boldsymbol{Q} \in \mathbb{C}^{m\times m}$*, we have*

$$\|\boldsymbol{QA}\|_2 = \|\boldsymbol{A}\|_2 \text{ and } \|\boldsymbol{QA}\|_F = \|\boldsymbol{A}\|_F.$$

In other words, 2-norm and Frobenius norms are invariant under unitary multiplication.

Proof for 2-norm: $\|\boldsymbol{Qy}\|_2 = \|\boldsymbol{y}\|_2$ for $\boldsymbol{y} \in \mathbb{C}^m$ and therefore $\|\boldsymbol{QAx}\|_2 = \|\boldsymbol{Ax}\|_2$ for $\boldsymbol{x} \in \mathbb{C}^n$. It then follows from definition of 2-norm.

Invariance under Unitary Multiplication

Theorem

For any $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ *and unitary* $\boldsymbol{Q} \in \mathbb{C}^{m \times m}$, *we have*

$$\|\boldsymbol{QA}\|_2 = \|\boldsymbol{A}\|_2 \text{ and } \|\boldsymbol{QA}\|_F = \|\boldsymbol{A}\|_F.$$

In other words, 2-norm and Frobenius norms are invariant under unitary multiplication.
Proof for 2-norm: $\|\boldsymbol{Qy}\|_2 = \|\boldsymbol{y}\|_2$ for $\boldsymbol{y} \in \mathbb{C}^m$ and therefore $\|\boldsymbol{QAx}\|_2 = \|\boldsymbol{Ax}\|_2$ for $\boldsymbol{x} \in \mathbb{C}^n$. It then follows from definition of 2-norm.
Proof for Frobenius norm:
$\|\boldsymbol{QA}\|_F = \operatorname{tr}((\boldsymbol{QA})^*\boldsymbol{QA}) = \operatorname{tr}(\boldsymbol{A}^*\boldsymbol{Q}^*\boldsymbol{QA}) = \operatorname{tr}(\boldsymbol{A}^*\boldsymbol{A}) = \|\boldsymbol{A}\|_F.$

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 $\boldsymbol{A}S$ 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 $\boldsymbol{AS}$
- $\boldsymbol{Av}_j = \sigma_j \boldsymbol{u}_j$ for $1 \leq j \leq n$

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 $\boldsymbol{A}\boldsymbol{v}_1/\sigma_1$ (or any unit-length vector if $\sigma_1 = 0$) and $\boldsymbol{v}_1$, respectively. Note that

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

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 therefore $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^*$ where $\boldsymbol{U} = \boldsymbol{U}_1 \begin{bmatrix} 1 & \boldsymbol{0}^* \\ \boldsymbol{0} & \boldsymbol{U}_2 \end{bmatrix}$, $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & \boldsymbol{0}^* \\ \boldsymbol{0} & \boldsymbol{\Sigma}_2 \end{bmatrix}$, and $\boldsymbol{V} = \boldsymbol{V}_1 \begin{bmatrix} 1 & \boldsymbol{0}^* \\ \boldsymbol{0} & \boldsymbol{V}_2 \end{bmatrix}$. $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary. □

Uniqueness of SVD

Theorem

*(Uniqueness) The singular values $\{\sigma_j\}$ are uniquely determined. If **A** is square and the σ_j are distinct, the left and right singular vectors are uniquely determined **up to complex signs**.*

Main idea of proof: 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. Once $\sigma_1, \boldsymbol{u}_1$, and $\boldsymbol{v}_1$ are determined, the reminder 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

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****.*

Main idea of proof: 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. Once $\sigma_1, \boldsymbol{u}_1$, and $\boldsymbol{v}_1$ are determined, the reminder 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 flip the sign of a singular vector?

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****.*

Main idea of proof: 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. Once $\sigma_1, \boldsymbol{u}_1$, and $\boldsymbol{v}_1$ are determined, the reminder 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 flip the sign of a singular vector?

- Question: What if σ_i is not distinct?

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

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

- **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.