AMS526: Numerical Analysis I (Numerical Linear Algebra)

Lecture 6: Projectors and QR Factorization

Xiangmin Jiao

SUNY Stony Brook

September 18, 2008

Outline

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.

SVD vs. Eigenvalue Decomposition

- Eigenvalue decomposition is

$$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$$

SVD vs. Eigenvalue Decomposition

- Eigenvalue decomposition is

$$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$$

- 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 nonnegative 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 is

$$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$$

- 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 nonnegative 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 $\boldsymbol{A}^*\boldsymbol{A}$ and $\boldsymbol{A}\boldsymbol{A}^*$
  - Singular values of hermitian matrices are absolute values of eigenvalues

Matrix Properties via SVD

- Analysis of rank, range, null space
- 2-norm and Frobenius norm
- Determinant of matrix

Matrix Properties via SVD

- Analysis of rank, range, null space
- 2-norm and Frobenius norm
- Determinant of matrix

- However, SVD may not be 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

Outline

Projectors

- A *projector* satisfies $\boldsymbol{P}^2 = \boldsymbol{P}$. They are also said to be *idempotent*.
  - *Orthogonal* projector
  - *Oblique* projector
- Complementary projectors: $\boldsymbol{P}$ vs. $\boldsymbol{I} - \boldsymbol{P}$.
- What space does $\boldsymbol{I} - \boldsymbol{P}$ project?
  - Answer: $\text{null}(\boldsymbol{P})$.
  - $\text{range}(\boldsymbol{I} - \boldsymbol{P}) \supseteq \text{null}(\boldsymbol{P})$ because $\boldsymbol{P}\boldsymbol{v} = \boldsymbol{0} \Rightarrow (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{v} = \boldsymbol{v}$.
  - $\text{range}(\boldsymbol{I} - \boldsymbol{P}) \subseteq \text{null}(\boldsymbol{P})$ because for any $\boldsymbol{v}$ $(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{v} = \boldsymbol{v} - \boldsymbol{P}\boldsymbol{v} \in \text{null}(\boldsymbol{P})$.
- A projector separates $\mathbb{C}^m$ into two complementary subspace: range space and null space.
- It projects onto the range space along the null space.

Orthogonal Projector

Definition

An *orthogonal projector* is one that projects onto a subspace S_1 along a space $\boldsymbol{S_2}$, where S_1 and S_2 are orthogonal.

Theorem

A projector $\boldsymbol{P}$ *is orthogonal if and only if* $\boldsymbol{P} = \boldsymbol{P}^*$.

Proof.

"If" direction: If $\boldsymbol{P} = \boldsymbol{P}^*$, then $(\boldsymbol{P}\boldsymbol{x})^*(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y} = \boldsymbol{x}(\boldsymbol{P} - \boldsymbol{P}^2)\boldsymbol{y} = 0$.
"Only if" direction: Use SVD. Suppose $\boldsymbol{P}$ projects onto S_1 along S_2 where$S_1 \perp S_2$, and S_1 has dimension n. Let $\{\boldsymbol{q}_1, \ldots, , \boldsymbol{q}_n\}$ be orthonormal basis of S_1 and $\{\boldsymbol{q}_{n+1}, \ldots, \boldsymbol{q}_m\}$ be a basis for S_2. Let $\boldsymbol{Q}$ be unitary matrix whose jth column is q_j, and we have

$$\boldsymbol{P}\boldsymbol{Q} = (\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n, \boldsymbol{0}, \ldots, \boldsymbol{0}),$$

so $\boldsymbol{Q}^*\boldsymbol{P}\boldsymbol{Q} = \mathrm{diag}(1, 1, \cdots, 1, 0, \cdots) = \boldsymbol{\Sigma}$, and $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{\Sigma}\boldsymbol{Q}^*$. □

Basis of Projections

- Projection with **orthonormal** basis
  - Given any matrix $\boldsymbol{Q}$ whose columns are orthonormal, then $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{Q}^*$ is orthogonal projector, so is $\boldsymbol{I} - \boldsymbol{P}$
  - We write $\boldsymbol{I} - \boldsymbol{P}$ as $\boldsymbol{P}_\perp$
  - In particular, if $\boldsymbol{Q} = \boldsymbol{q}$, we write $\boldsymbol{P}_{\boldsymbol{q}} = \boldsymbol{q}\boldsymbol{q}^*$ and $\boldsymbol{P}_{\perp \boldsymbol{q}} = \boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{q}}$

Basis of Projections

- Projection with **orthonormal** basis
  - Given any matrix $\boldsymbol{Q}$ whose columns are orthonormal, then $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{Q}^*$ is orthogonal projector, so is $\boldsymbol{I} - \boldsymbol{P}$
  - We write $\boldsymbol{I} - \boldsymbol{P}$ as $\boldsymbol{P}_\perp$
  - In particular, if $\boldsymbol{Q} = \boldsymbol{q}$, we write $\boldsymbol{P}_{\boldsymbol{q}} = \boldsymbol{q}\boldsymbol{q}^*$ and $\boldsymbol{P}_{\perp\boldsymbol{q}} = \boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{q}}$
- Projection with **arbitrary** basis
  - Given any matrix $\boldsymbol{A}$ that has full rank

$$\boldsymbol{P} = \boldsymbol{A}(\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*$$

    is an orthogonal projection, which is easy to verify
  - What does $\boldsymbol{P}$ project onto? range($\boldsymbol{A}$)

Outline

Motivation

Question: Given a linear system $\boldsymbol{Ax} = \boldsymbol{b}$ where $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ $(m \geq n)$ with full rank, how to solve the linear system?

Motivation

Question: Given a linear system $\boldsymbol{Ax} = \boldsymbol{b}$ where $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ ($m \geq n$) with full rank, how to solve the linear system?
Answer: One possible solution is to use SVD. How?

Motivation

Question: Given a linear system $\boldsymbol{Ax} = \boldsymbol{b}$ where $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ $(m \geq n)$ with full rank, how to solve the linear system?
Answer: One possible solution is to use SVD. How?

$$\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V}^*, \quad \text{so } \boldsymbol{x} = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^*\boldsymbol{b}.$$

Motivation

Question: Given a linear system $\boldsymbol{Ax} = \boldsymbol{b}$ where $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ $(m \geq n)$ with full rank, how to solve the linear system?
Answer: One possible solution is to use SVD. How?

$$\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V}^*, \quad \text{so } \boldsymbol{x} = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^*\boldsymbol{b}.$$

Another solution is to use QR factorization, which decompose $\boldsymbol{A}$ into product of two simple matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ where columns of $\boldsymbol{Q}$ are orthonormal and $\boldsymbol{R}$ is upper triangular.

Two Different Versions of QR

Analogous to SVD, there are two versions of QR

- Full QR factorization: $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ $(m \geq n)$

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$$

  where $\boldsymbol{Q} \in \mathbb{C}^{m \times m}$ is unitary and $\boldsymbol{R} \in \mathbb{C}^{m \times n}$ is upper triangular

- Reduced QR factorization: $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ $(m \geq n)$

$$\boldsymbol{A} = \widehat{\boldsymbol{Q}}\widehat{\boldsymbol{R}}$$

  where $\boldsymbol{Q} \in \mathbb{C}^{m \times n}$ contains orthonormal vectors and $\boldsymbol{R} \in \mathbb{C}^{n \times n}$ is upper triangular

- Question: What space do column vectors $\{\boldsymbol{q}_j\}$, $n < j \leq m$ span?

Two Different Versions of QR

Analogous to SVD, there are two versions of QR

- Full QR factorization: $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ $(m \geq n)$

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$$

  where $\boldsymbol{Q} \in \mathbb{C}^{m \times m}$ is unitary and $\boldsymbol{R} \in \mathbb{C}^{m \times n}$ is upper triangular

- Reduced QR factorization: $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ $(m \geq n)$

$$\boldsymbol{A} = \widehat{\boldsymbol{Q}}\widehat{\boldsymbol{R}}$$

  where $\boldsymbol{Q} \in \mathbb{C}^{m \times n}$ contains orthonormal vectors and $\boldsymbol{R} \in \mathbb{C}^{n \times n}$ is upper triangular

- Question: What space do column vectors $\{\boldsymbol{q}_j\}$, $n < j \leq m$ span?

- Answer: First j column vectors of $\boldsymbol{A}$, i.e., $\langle \boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_j \rangle = \langle \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_j \rangle$.

Gram-Schmidt Orthogonalization

- A method to construct QR factorization is to orthogonalize the column vectors of $\boldsymbol{A}$:
- Basic idea:
  - Take first column $\boldsymbol{a}_1$ and normalize it to obtain vector $\boldsymbol{q}_1$;
  - Take second column $\boldsymbol{a}_2$, subtract its orthogonal projection to $\boldsymbol{q}_1$, and normalize to obtain $\boldsymbol{q}_2$;
  - ...
  - Take jth column of $\boldsymbol{a}_j$, subtract its orthogonal projection to $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{j-1}$, and normalize to obtain $\boldsymbol{q}_j$;

$$\boldsymbol{v}_j = \boldsymbol{a}_j - \sum_{i=1}^{j-1} \boldsymbol{q}_i^* \boldsymbol{a}_j \boldsymbol{q}_i, \;\; \boldsymbol{q}_j = \boldsymbol{v}_j / \|\boldsymbol{v}_j\|.$$

- This idea is called *Gram-Schmidt orthogonalization*.