Lecture 05
Projectors and QR Factorization

Songting Luo

Department of Mathematics
Iowa State University

MATH 562 Numerical Analysis II

Outline

1. Projectors

2. QR Factorization

Outline

1. Projectors

2. QR Factorization

Projectors

- A projector satisfies $\mathbf{P}^2 = \mathbf{P}$. They are also said to be idempotent.
  - Orthogonal projector. (Projects onto S_1 along S_2 where S_2 is orthogonal to S_1).
  - Oblique projector. (non-orthogonal)
- Example: $\begin{bmatrix} 0 & 0 \\ \alpha & 1 \end{bmatrix}$
  - is an oblique projector if $\alpha \neq 0$
  - is orthogonal projector if $\alpha = 0$.

Complementary Projectors

- Complementary projector to $\mathbf{P}$: $\mathbf{I}-\mathbf{P}$.
- What space does $\mathbf{I}-\mathbf{P}$ project?
  - What is the range? What is the null space?

Complementary Projectors

- Complementary projector to $\mathbf{P}$: $\mathbf{I} - \mathbf{P}$.
- What space does $\mathbf{I} - \mathbf{P}$ project?
  - What is the range? What is the null space?
  - Answer: $null(\mathbf{P}) = range(\mathbf{I} - \mathbf{P})$
  - Also, By $\mathbf{P} = \mathbf{I} - (\mathbf{I} - \mathbf{P})$, $null(\mathbf{I} - \mathbf{P}) = range(\mathbf{P})$.

Complementary Projectors $\mathbf{P}$ vs. $\mathbf{I} - \mathbf{P}$

- A projector separates $\mathbb{C}^m$ into two complementary subspaces: range space and null space. For projector $\mathbf{P} \in \mathbb{C}^{m \times m}$,
  - $range(\mathbf{P}) + null(\mathbf{P}) = \mathbb{C}^m$
  - $range(\mathbf{P}) \bigcap null(\mathbf{P}) = \{\mathbf{0}\}$
  - It projects onto range space along null space. In other words,

$$\mathbf{x} = \mathbf{P}\mathbf{x} + \mathbf{r}, \quad \text{where } \mathbf{r} \in null(\mathbf{P})$$

Complementary Projectors $\mathbf{P}$ vs. $\mathbf{I} - \mathbf{P}$

- A projector separates $\mathbb{C}^m$ into two complementary subspaces: range space and null space. For projector $\mathbf{P} \in \mathbb{C}^{m \times m}$,
  - $range(\mathbf{P}) + null(\mathbf{P}) = \mathbb{C}^m$
  - $range(\mathbf{P}) \bigcap null(\mathbf{P}) = \{\mathbf{0}\}$
  - It projects onto range space along null space. In other words,

$$\mathbf{x} = \mathbf{P}\mathbf{x} + \mathbf{r}, \quad \text{where } \mathbf{r} \in null(\mathbf{P})$$

- Conversely, given two complementary subspaces S_1 and S_2, there is a projector $\mathbf{P}$ such that $rang(\mathbf{P}) = S_1$, and $null(\mathbf{P}) = S_2$. (Say $\mathbf{P}$ is the projector onto S_1 along S_2.)
- **Question:** Are range space and null space of projector orthogonal to each other?

Orthogonal Projector

- An orthogonal projector is one that projects onto a subspace S_1 along a space S_2, where S_1 and S_2 are orthogonal.

Theorem (Existence)

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

Orthogonal Projector

- An orthogonal projector is one that projects onto a subspace S_1 along a space S_2, where S_1 and S_2 are orthogonal.

Theorem (Existence)

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

Proof

"If" direction: if $\mathbf{P} = \mathbf{P}^*$, then $(\mathbf{P}\mathbf{x})^*(\mathbf{I} - \mathbf{P})\mathbf{y} = \mathbf{x}^*(\mathbf{P} - \mathbf{P}^2)\mathbf{y} = 0$

"Only if" direction: use SVD, Suppose $\mathbf{P}$ projects onto S_1 along S_2 where $S_1 \perp S_2$, and S_1 has dimension n. Let $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ be orthonormal basis of S_1 and $\{\mathbf{q}_{n+1}, \ldots, \mathbf{q}_m\}$ be a basis for S_2. Let $\mathbf{Q}$ be unitary matrix whose jth column is $\mathbf{q}_j$, and we have $\mathbf{PQ} = [\mathbf{q}_1, \ldots, \mathbf{q}_n, \mathbf{0}, \ldots, \mathbf{0}]$, so $\mathbf{Q}^*\mathbf{PA} = diag(1, 1, \cdots, 0, \cdots) = \mathbf{\Sigma}$, and $\mathbf{P} = \mathbf{Q\Sigma Q}^*$.

Orthogonal Projector

- An orthogonal projector is one that projects onto a subspace S_1 along a space S_2, where S_1 and S_2 are orthogonal.

Theorem (Existence)

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

Proof

"If" direction: if $\mathbf{P} = \mathbf{P}^*$, then $(\mathbf{Px})^*(\mathbf{I} - \mathbf{P})\mathbf{y} = \mathbf{x}^*(\mathbf{P} - \mathbf{P}^2)\mathbf{y} = 0$

"Only if" direction: use SVD, Suppose $\mathbf{P}$ projects onto S_1 along S_2 where $S_1 \perp S_2$, and S_1 has dimension n. Let $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$ be orthonormal basis of S_1 and $\{\mathbf{q}_{n+1}, \dots, \mathbf{q}_m\}$ be a basis for S_2. Let $\mathbf{Q}$ be unitary matrix whose jth column is $\mathbf{q}_j$, and we have $\mathbf{PQ} = [\mathbf{q}_1, \dots, \mathbf{q}_n, \mathbf{0}, \dots, \mathbf{0}]$, so $\mathbf{Q}^*\mathbf{PA} = diag(1, 1, \cdots, 0, \cdots) = \mathbf{\Sigma}$, and $\mathbf{P} = \mathbf{Q\Sigma Q}^*$.

Question: Are orthogonal projectors orthogonal matrices?

Basis of Projections

- Projection with orthonormal basis
  - Given any matrix $\hat{\mathbf{Q}} \in \mathbb{C}^{m \times n}$ whose columns are orthonormal, then $\mathbf{P} = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^*$ is orthogonal projector, so is $\mathbf{I} - \mathbf{P}$.
  - We write $\mathbf{I} - \mathbf{P}$ as $\mathbf{P}_\perp$
  - In particular, if $\hat{\mathbf{Q}} = \mathbf{q}$, we write $\mathbf{P_q} = \mathbf{qq}^*$ and $\mathbf{P}_{\perp \mathbf{q}} = \mathbf{I} - \mathbf{P_q}$.
  - For arbitrary vector $\mathbf{a}$, we write $\mathbf{P_a} = \frac{\mathbf{aa}^*}{\mathbf{a}^*\mathbf{a}}$ and $\mathbf{P}_{\perp \mathbf{a}} = \mathbf{I} - \mathbf{P_a}$.

Basis of Projections

- Projection with orthonormal basis
  - Given any matrix $\hat{\mathbf{Q}} \in \mathbb{C}^{m \times n}$ whose columns are orthonormal, then $\mathbf{P} = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^*$ is orthogonal projector, so is $\mathbf{I} - \mathbf{P}$.
  - We write $\mathbf{I} - \mathbf{P}$ as $\mathbf{P}_\perp$
  - In particular, if $\hat{\mathbf{Q}} = \mathbf{q}$, we write $\mathbf{P}_\mathbf{q} = \mathbf{q}\mathbf{q}^*$ and $\mathbf{P}_{\perp\mathbf{q}} = \mathbf{I} - \mathbf{P}_\mathbf{q}$.
  - For arbitrary vector $\mathbf{a}$, we write $\mathbf{P}_\mathbf{a} = \frac{\mathbf{a}\mathbf{a}^*}{\mathbf{a}^*\mathbf{a}}$ and $\mathbf{P}_{\perp\mathbf{a}} = \mathbf{I} - \mathbf{P}_\mathbf{a}$.
- Projection with arbitrary basis
  - Given any matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ that has full rank and $m \geqslant n$, then

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

    is orthogonal projection
  - What does $\mathbf{P}$ project onto?

Basis of Projections

- Projection with orthonormal basis
  - Given any matrix $\hat{\mathbf{Q}} \in \mathbb{C}^{m \times n}$ whose columns are orthonormal, then $\mathbf{P} = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^*$ is orthogonal projector, so is $\mathbf{I} - \mathbf{P}$.
  - We write $\mathbf{I} - \mathbf{P}$ as $\mathbf{P}_\perp$
  - In particular, if $\hat{\mathbf{Q}} = \mathbf{q}$, we write $\mathbf{P_q} = \mathbf{qq}^*$ and $\mathbf{P}_{\perp\mathbf{q}} = \mathbf{I} - \mathbf{P_q}$.
  - For arbitrary vector $\mathbf{a}$, we write $\mathbf{P_a} = \frac{\mathbf{aa}^*}{\mathbf{a}^*\mathbf{a}}$ and $\mathbf{P}_{\perp\mathbf{a}} = \mathbf{I} - \mathbf{P_a}$.
- Projection with arbitrary basis
  - Given any matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ that has full rank and $m \geqslant n$, then

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

    is orthogonal projection
  - What does $\mathbf{P}$ project onto? Answer: $range(\mathbf{A})$.
  - $(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*$ is called the pseudo-inverse of $\mathbf{A}$, denoted as $\mathbf{A}^+$.

Basis of Projections

- Projection with orthonormal basis
  - Given any matrix $\hat{\mathbf{Q}} \in \mathbb{C}^{m \times n}$ whose columns are orthonormal, then $\mathbf{P} = \hat{\mathbf{Q}}\hat{\mathbf{Q}}^*$ is orthogonal projector, so is $\mathbf{I} - \mathbf{P}$.
  - We write $\mathbf{I} - \mathbf{P}$ as $\mathbf{P}_\perp$
  - In particular, if $\hat{\mathbf{Q}} = \mathbf{q}$, we write $\mathbf{P_q} = \mathbf{qq}^*$ and $\mathbf{P}_{\perp\mathbf{q}} = \mathbf{I} - \mathbf{P_q}$.
  - For arbitrary vector $\mathbf{a}$, we write $\mathbf{P_a} = \frac{\mathbf{aa}^*}{\mathbf{a}^*\mathbf{a}}$ and $\mathbf{P}_{\perp\mathbf{a}} = \mathbf{I} - \mathbf{P_a}$.
- Projection with arbitrary basis
  - Given any matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ that has full rank and $m \geqslant n$, then

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

    is orthogonal projection
  - What does $\mathbf{P}$ project onto? Answer: $range(\mathbf{A})$.
  - $(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*$ is called the pseudo-inverse of $\mathbf{A}$, denoted as $\mathbf{A}^+$.

Outline

1. Projectors

2. QR Factorization

Motivation

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

Motivation

Question: Given a linear system $\mathbf{A}\mathbf{x} \approx \mathbf{b}$ where $\mathbf{A} \in \mathbb{C}^{m \times n}$ $(m \geqslant n)$ has full rank, how to solve the linear system?
Answer: One possible solution is to use SVD. How?

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

Motivation

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

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

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

Two Different Versions of QR

Analogous to SVD, there are two versions of QR

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

$$\mathbf{A} = \mathbf{QR}$$

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

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

$$\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$$

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

- What space do $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_j\},\ j \leqslant n$ span?

Two Different Versions of QR

Analogous to SVD, there are two versions of QR

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

$$\mathbf{A} = \mathbf{QR}$$

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

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

$$\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$$

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

- What space do $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_j\}$, $j \leqslant n$ span?
  - Answer: for full rank $\mathbf{A}$, first j columns of $\mathbf{A}$, i.e.,

$$\langle \mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_j \rangle = \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_j \rangle$$

Gram-Schmidt Orthogonalization

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

$$\mathbf{v}_j = \mathbf{a}_j - \sum_{i=1}^{j-1} \mathbf{q}_i^* \mathbf{a}_i \mathbf{q}_i, \quad \mathbf{q}_j = \mathbf{v}_j / \|\mathbf{v}_j\|_2.$$

- This idea is called Gram-Schmidt orthogonalization.

Gram-Schmidt Projections

- Orthogonal vectors produced by Gram-Schmidt can be written in terms of projectors

$$\mathbf{q}_j = \frac{\mathbf{P}_j \mathbf{a}_j}{\|\mathbf{P}_j \mathbf{a}_j\|_2}$$

where

$$\mathbf{P}_j = \mathbf{I} - \hat{\mathbf{Q}}_{j-1}\hat{\mathbf{Q}}_{j-1}^*, \text{ with } \hat{\mathbf{Q}}_{j-1} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_{j-1}]$$

- $\mathbf{P}_j$ projects orthogonally onto space orthogonal to $\langle \mathbf{q}_1, \ldots, \mathbf{q}_{j-1} \rangle$ and rank of $\mathbf{P}_j$ is $m - (j-1)$.

Algorithm of Gram-Schmidt Orthogonalization

Classical Gram-Schmidt method

for $j = 1$ to n

$\quad \mathbf{v}_j = \mathbf{a}_j$

$\quad$ for $i = 1$ to $j-1$

$\quad\quad r_{ij} = \mathbf{q}_i^* \mathbf{a}_j$

$\quad\quad \mathbf{v}_j = \mathbf{v}_j - r_{ij}\mathbf{q}_i$

$\quad r_{ij} = \|\mathbf{v}_j\|_2$

$\quad \mathbf{q}_j = \mathbf{v}_j / r_{ij}$

Algorithm of Gram-Schmidt Orthogonalization

Classical Gram-Schmidt method

for $j = 1$ to n
  $\mathbf{v}_j = \mathbf{a}_j$
  for $i = 1$ to $j-1$
    $r_{ij} = \mathbf{q}_i^* \mathbf{a}_j$
    $\mathbf{v}_j = \mathbf{v}_j - r_{ij}\mathbf{q}_i$
  $r_{ij} = \|\mathbf{v}_j\|_2$
  $\mathbf{q}_j = \mathbf{v}_j / r_{ij}$

- Classical Gram-Schmidt (CGS) is unstable, which means that its solution is sensitive to perturbation

Existence of QR

Theorem

Every $\mathbf{A} \in \mathbb{C}^{m\times n}(m \geqslant n)$ *has full QR factorization, hence also a reduced QR factorization.*

Existence of QR

Theorem

Every $\mathbf{A} \in \mathbb{C}^{m \times n} (m \geqslant n)$ *has full QR factorization, hence also a reduced QR factorization.*

Key idea of proof

If $\mathbf{A}$ has full rank, Gram-Schmidt algorithm provides a proof itself for having reduced QR.
If $\mathbf{A}$ does not have full rank, at some step $\mathbf{v}_j = \mathbf{0}$. We can set $\mathbf{q}_j$ to be a vector orthogonal to $\mathbf{q}_i,\ i < j$.
To construct full QR from reduced QR, just continue Gram-Schmidt algorithm additional $m - n$ steps.

Uniqueness of QR

Theorem

Every $\mathbf{A} \in \mathbb{C}^{m\times n}(m \geqslant n)$ *of full rank has a unique reduced QR factorization* $\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ *with* $r_{ij} > 0$.

Key idea of proof

Proof is provided by Gram-Schmidt iteration itself. If the signs of r_{jj} are determined, then r_{ij} and $\mathbf{q}_j$ are determined.