

Lecture 3: Projector, Classical/Modified Gram–Schmidt orthogonalization, QR factorization



School of Mathematical Sciences, Xiamen University

1. Projector

- A square matrix $\mathbf{P} \in \mathbb{C}^{m \times m}$ is called a *projector* if $\mathbf{P}^2 = \mathbf{P}$. Any projector is diagonalizable. (Eigenvalues?) **Example:** $\mathbf{P} = \begin{bmatrix} 0 & 0 \\ \alpha & 1 \end{bmatrix}$

Theorem 1

Let \mathbf{P} be a projector. Then,

- (1) for all $\mathbf{v} \in \text{range}(\mathbf{P})$, we have $\mathbf{P}\mathbf{v} = \mathbf{v}$;
- (2) $\text{range}(\mathbf{P})$ and $\text{null}(\mathbf{P})$ satisfy

$$\text{range}(\mathbf{P}) \cap \text{null}(\mathbf{P}) = \{\mathbf{0}\}, \quad \text{range}(\mathbf{P}) + \text{null}(\mathbf{P}) = \mathbb{C}^m;$$

- (3) $\mathbf{I} - \mathbf{P}$ is a projector, and

$$\text{range}(\mathbf{I} - \mathbf{P}) = \text{null}(\mathbf{P}), \quad \text{null}(\mathbf{I} - \mathbf{P}) = \text{range}(\mathbf{P}).$$

- (4) if $\mathbf{P} \neq \mathbf{0}, \mathbf{I}$, we have $\|\mathbf{I} - \mathbf{P}\|_2 = \|\mathbf{P}\|_2$. (See [Ref. 1](#) and [Ref. 2](#))

- Two subspaces $\mathcal{S}_1, \mathcal{S}_2 \subseteq \mathbb{C}^m$ are called *complementary subspaces* if they satisfy

$$\mathcal{S}_1 \cap \mathcal{S}_2 = \{\mathbf{0}\}, \quad \mathcal{S}_1 + \mathcal{S}_2 = \mathbb{C}^m.$$

Theorem 2

Let \mathcal{S}_1 and \mathcal{S}_2 be complementary subspaces. Then there exists a unique projector \mathbf{P} with $\text{range}(\mathbf{P}) = \mathcal{S}_1$ and $\text{null}(\mathbf{P}) = \mathcal{S}_2$.

Proof.

The existence is left as an exercise. Now we prove the uniqueness. Let \mathbf{e}_j denote the j th column of the identity matrix \mathbf{I} . Since \mathcal{S}_1 and \mathcal{S}_2 are complementary, we can assume $\mathbf{e}_j = \mathbf{s}_j^1 + \mathbf{s}_j^2$, where $\mathbf{s}_j^1 \in \mathcal{S}_1$, and $\mathbf{s}_j^2 \in \mathcal{S}_2$. Assume both \mathbf{P}_1 and \mathbf{P}_2 are desired projectors. Then we have

$$\begin{aligned} \forall 1 \leq j \leq m, \quad (\mathbf{P}_1 - \mathbf{P}_2)\mathbf{e}_j &= (\mathbf{P}_1 - \mathbf{P}_2)\mathbf{s}_j^1 + (\mathbf{P}_1 - \mathbf{P}_2)\mathbf{s}_j^2 \\ &= \mathbf{P}_1\mathbf{s}_j^1 - \mathbf{P}_2\mathbf{s}_j^1 = \mathbf{s}_j^1 - \mathbf{s}_j^1 = \mathbf{0}. \end{aligned}$$

Therefore, $\mathbf{P}_1 = \mathbf{P}_2$, i.e., uniqueness. □

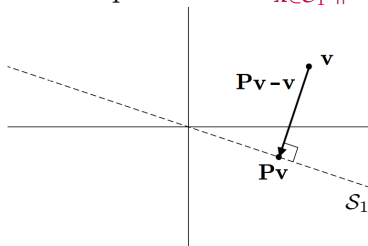
- Let \mathcal{S}_1 and \mathcal{S}_2 be complementary subspaces. The unique projector \mathbf{P} with $\text{range}(\mathbf{P}) = \mathcal{S}_1$ and $\text{null}(\mathbf{P}) = \mathcal{S}_2$ is called the *projector onto \mathcal{S}_1 along \mathcal{S}_2* .

1.1. Orthogonal and oblique projectors

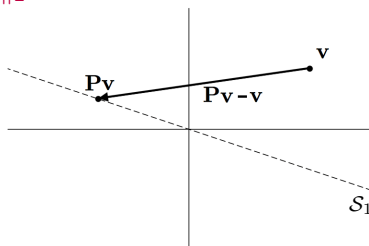
- For a projector \mathbf{P} , if $\text{range}(\mathbf{P})$ and $\text{null}(\mathbf{P})$ are orthogonal, then it is called an *orthogonal* projector. Otherwise, *oblique*.

Warning: orthogonal projector \neq orthogonal matrix!!!

- Geometric interpretation: consider projector \mathbf{P} s.t. $\text{range}(\mathbf{P}) = \mathcal{S}_1$ and the problem $\min_{\mathbf{x} \in \mathcal{S}_1} \|\mathbf{v} - \mathbf{x}\|_2$.



The orthogonal projection



An oblique projection

Theorem 3

A matrix \mathbf{P} is an orthogonal projector if and only if it is idempotent ($\mathbf{P}^2 = \mathbf{P}$) and Hermitian ($\mathbf{P} = \mathbf{P}^*$).

- $\mathbf{P} = \begin{bmatrix} 0 & 0 \\ \alpha & 1 \end{bmatrix}$: oblique (if $\alpha \neq 0$) or orthogonal (if $\alpha = 0$) projector.

Theorem 4

Let the columns of \mathbf{Q}_r be an orthonormal basis of an r -dimensional subspace \mathcal{S} . Then the orthogonal projector onto \mathcal{S} is given by $\mathbf{Q}_r \mathbf{Q}_r^*$, and the orthogonal projector onto \mathcal{S}^\perp is given by $\mathbf{I} - \mathbf{Q}_r \mathbf{Q}_r^*$.

- $\mathbf{a} \neq \mathbf{0}$, $\mathbf{P}_{\mathbf{a}} = \frac{\mathbf{a}\mathbf{a}^*}{\mathbf{a}^*\mathbf{a}}$ onto $\text{span}\{\mathbf{a}\}$, $\mathbf{P}_{\mathbf{a}^\perp} = \mathbf{I} - \frac{\mathbf{a}\mathbf{a}^*}{\mathbf{a}^*\mathbf{a}}$ onto $\text{span}\{\mathbf{a}\}^\perp$
- Let $\mathbf{A} \in \mathbb{C}^{m \times n}$. The orthogonal projector onto $\text{range}(\mathbf{A})$ is given by $\mathbf{U}_r \mathbf{U}_r^*$, where \mathbf{U}_r is the matrix in SVD of \mathbf{A} .
- Others: $\mathbf{A}\mathbf{A}^\dagger$ onto $\text{range}(\mathbf{A})$, $\mathbf{A}^\dagger\mathbf{A}$ onto $\text{range}(\mathbf{A}^*)$

1.2. Distance between subspaces and CS decomposition

Definition 5

Let $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{C}^m$ be two subspaces with $\dim(\mathcal{X}) = \dim(\mathcal{Y})$. Let $\mathbf{P}_{\mathcal{X}}$ and $\mathbf{P}_{\mathcal{Y}}$ be the orthogonal projectors onto \mathcal{X} and \mathcal{Y} , respectively. The distance between \mathcal{X} and \mathcal{Y} is defined as

$$\text{dist}(\mathcal{X}, \mathcal{Y}) = \|\mathbf{P}_{\mathcal{X}} - \mathbf{P}_{\mathcal{Y}}\|_2.$$

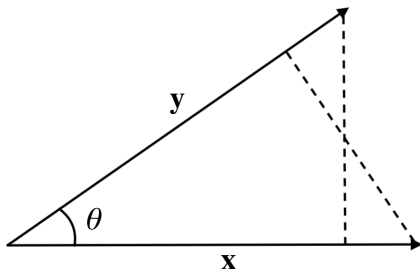
- **Example:** Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ with $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$ and $\mathbf{x} \neq \mathbf{y}$. By

$$\begin{aligned} \mathbf{x}\mathbf{x}^\top - \mathbf{y}\mathbf{y}^\top &= \mathbf{x}(\mathbf{x} - \mathbf{y}^\top \mathbf{x} \mathbf{y})^\top + (\mathbf{x}^\top \mathbf{y} \mathbf{x} - \mathbf{y})\mathbf{y}^\top \\ &= \begin{bmatrix} \mathbf{x} & \frac{\mathbf{x}^\top \mathbf{y} \mathbf{x} - \mathbf{y}}{\|\mathbf{x}^\top \mathbf{y} \mathbf{x} - \mathbf{y}\|_2} \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} \frac{\mathbf{x} - \mathbf{y}^\top \mathbf{x} \mathbf{y}}{\|\mathbf{x} - \mathbf{y}^\top \mathbf{x} \mathbf{y}\|_2} & \mathbf{y} \end{bmatrix}^\top \end{aligned}$$

with $\sigma_1 = \|\mathbf{x} - \mathbf{y}^\top \mathbf{x} \mathbf{y}\|_2$ and $\sigma_2 = \|\mathbf{x}^\top \mathbf{y} \mathbf{x} - \mathbf{y}\|_2$, we have

$$\begin{aligned} \text{dist}(\text{span}\{\mathbf{x}\}, \text{span}\{\mathbf{y}\}) &= \left\| \mathbf{x}\mathbf{x}^\top - \mathbf{y}\mathbf{y}^\top \right\|_2 = \sigma_1 = \sigma_2 \\ &= \sqrt{1 - |\mathbf{x}^\top \mathbf{y}|^2} = \sin \theta. \end{aligned}$$

- Geometric interpretation for the case $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2, \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$



The distance between $\text{span}\{\mathbf{x}\}$ and $\text{span}\{\mathbf{y}\}$ is

$$\text{dist}(\text{span}\{\mathbf{x}\}, \text{span}\{\mathbf{y}\}) = \sqrt{1 - |\mathbf{x}^\top \mathbf{y}|^2} = \sin \theta.$$

- Can this result be generalized to higher dimensional subspaces?
Read Pages 33–41 of [Numerical Linear Algebra by Zhihao Cao](#).

Theorem 6 (CS decomposition of unitary matrix)

Let

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \in \mathbb{C}^{m \times m}$$

be unitary, where $\mathbf{Q}_{11} \in \mathbb{C}^{r \times r}$, $\mathbf{Q}_{12} \in \mathbb{C}^{r \times (m-r)}$, $\mathbf{Q}_{21} \in \mathbb{C}^{(m-r) \times r}$, and $\mathbf{Q}_{22} \in \mathbb{C}^{(m-r) \times (m-r)}$. Assume that $r \leq m/2$. Then there exist unitary matrices $\mathbf{U}_1, \mathbf{V}_1 \in \mathbb{C}^{r \times r}$, and $\mathbf{U}_2, \mathbf{V}_2 \in \mathbb{C}^{(m-r) \times (m-r)}$ such that

$$\begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_1 & \\ & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{C} & -\mathbf{S} & \mathbf{0} \\ \mathbf{S} & \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1 & \\ & \mathbf{V}_2 \end{bmatrix}^*,$$

where

$$\mathbf{C} = \text{diag}\{c_1, \dots, c_r\}, \quad \mathbf{S} = \text{diag}\{s_1, \dots, s_r\}$$

with

$$c_i = \cos \theta_i, \quad s_i = \sin \theta_i, \quad \frac{\pi}{2} \geq \theta_1 \geq \dots \geq \theta_r \geq 0.$$

Theorem 7

Let \mathcal{X} and \mathcal{Y} be two r -dimensional subspaces of \mathbb{C}^m . Let the columns of \mathbf{X}_r and \mathbf{Y}_r be orthonormal bases of \mathcal{X} and \mathcal{Y} , respectively. Then,

$$\text{dist}(\mathcal{X}, \mathcal{Y}) = \sqrt{1 - \sigma_{\min}^2(\mathbf{X}_r^* \mathbf{Y}_r)},$$

where $\sigma_{\min}(\cdot)$ is the smallest singular value.

Proposition 8

Let $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{C}^m$ be two subspaces with $\dim(\mathcal{X}) \neq \dim(\mathcal{Y})$. Let $\mathbf{P}_{\mathcal{X}}$ and $\mathbf{P}_{\mathcal{Y}}$ be the orthogonal projectors onto \mathcal{X} and \mathcal{Y} , respectively. We have

$$\|\mathbf{P}_{\mathcal{X}} - \mathbf{P}_{\mathcal{Y}}\|_2 = 1.$$

Hint:

By $(\mathbf{P}_{\mathcal{X}} - \mathbf{P}_{\mathcal{Y}})^2 + (\mathbf{I} - \mathbf{P}_{\mathcal{X}} - \mathbf{P}_{\mathcal{Y}})^2 = \mathbf{I}$, we can show $\|\mathbf{P}_{\mathcal{X}} - \mathbf{P}_{\mathcal{Y}}\|_2 \leq 1$.
By $\exists \mathbf{x} (\neq \mathbf{0}) \in \{\mathcal{X} \cap \mathcal{Y}^\perp \text{ or } \mathcal{X}^\perp \cap \mathcal{Y}\}$, we can show $\|\mathbf{P}_{\mathcal{X}} - \mathbf{P}_{\mathcal{Y}}\|_2 \geq 1$. \square

1.3. General definitions

- Suppose that $\langle \cdot, \cdot \rangle$ denotes an inner product on a linear space \mathbb{V} . A linear mapping $\mathbf{T} : \mathbb{V} \mapsto \mathbb{V}$ is called
 - *idempotent* if for all $\mathbf{x} \in \mathbb{V}$, $\mathbf{T}(\mathbf{T}\mathbf{x}) = \mathbf{T}\mathbf{x}$;
 - *self-adjoint* (with respect to $\langle \cdot, \cdot \rangle$) if for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$,

$$\langle \mathbf{T}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{T}\mathbf{y} \rangle;$$

- an *orthogonal projector* (with respect to $\langle \cdot, \cdot \rangle$) if for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$,

$$\langle \mathbf{x} - \mathbf{T}\mathbf{x}, \mathbf{T}\mathbf{y} \rangle = 0.$$

- **Exercise:** Prove that if \mathbf{T} is self-adjoint, so is $\mathbf{I} - \mathbf{T}$ and vice versa.
- **Exercise:** Prove that for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$,

$$\langle \mathbf{x} - \mathbf{T}\mathbf{x}, \mathbf{T}\mathbf{y} \rangle = 0 \Leftrightarrow \mathbf{T}(\mathbf{T}\mathbf{x}) = \mathbf{T}\mathbf{x} \text{ and } \langle \mathbf{T}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{T}\mathbf{y} \rangle.$$

This means that

orthogonal projector \Leftrightarrow idempotent + self-adjoint.

2. Gram–Schmidt orthogonalization (GS)

- For n linearly independent vectors $\{\mathbf{a}_i\}_{i=1}^n$: at the j th step, Gram–Schmidt orthogonalization finds a unit vector \mathbf{q}_j that is orthogonal to $\mathbf{q}_1, \dots, \mathbf{q}_{j-1}$, lies in $\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_j\}$ as follows:

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

More generally, for a given inner product $\langle \cdot, \cdot \rangle$,

$$\tilde{\mathbf{q}}_j = \mathbf{a}_j - \sum_{i=1}^{j-1} \langle \mathbf{a}_j, \mathbf{q}_i \rangle \mathbf{q}_i, \quad \mathbf{q}_j = \frac{\tilde{\mathbf{q}}_j}{\sqrt{\langle \tilde{\mathbf{q}}_j, \tilde{\mathbf{q}}_j \rangle}}.$$

- Gram–Schmidt orthogonalization can also be represented via orthogonal projectors. For the standard inner product, we have

$$\tilde{\mathbf{q}}_j = \mathbf{P}_j \mathbf{a}_j, \quad \mathbf{q}_j = \tilde{\mathbf{q}}_j / \|\tilde{\mathbf{q}}_j\|_2,$$

where $\mathbf{P}_j = \mathbf{I} - \mathbf{Q}_{j-1} \mathbf{Q}_{j-1}^*$ and $\mathbf{Q}_{j-1} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \dots \ \mathbf{q}_{j-1}]$.

2.1. Classical Gram–Schmidt orthogonalization (CGS)

- CGS is based on the use of

$$\begin{aligned}\tilde{\mathbf{q}}_j &= \mathbf{P}_j \mathbf{a}_j = (\mathbf{I} - \mathbf{q}_1 \mathbf{q}_1^* - \mathbf{q}_2 \mathbf{q}_2^* - \cdots - \mathbf{q}_{j-1} \mathbf{q}_{j-1}^*) \mathbf{a}_j \\ &= \mathbf{a}_j - \mathbf{q}_1^* \mathbf{a}_j \mathbf{q}_1 - \mathbf{q}_2^* \mathbf{a}_j \mathbf{q}_2 \cdots - \mathbf{q}_{j-1}^* \mathbf{a}_j \mathbf{q}_{j-1}\end{aligned}$$

and calculates \mathbf{q}_j by evaluating the following formulas in order:

$$\begin{aligned}\mathbf{q}_j^{(0)} &= \mathbf{a}_j, \\ \mathbf{q}_j^{(1)} &= \mathbf{q}_j^{(0)} - \mathbf{q}_1^* \mathbf{a}_j \mathbf{q}_1, & r_{1j} &= \mathbf{q}_1^* \mathbf{a}_j, \\ \mathbf{q}_j^{(2)} &= \mathbf{q}_j^{(1)} - \mathbf{q}_2^* \mathbf{a}_j \mathbf{q}_2, & r_{2j} &= \mathbf{q}_2^* \mathbf{a}_j, \\ &\vdots & &\vdots \\ \mathbf{q}_j^{(j-1)} &= \mathbf{q}_j^{(j-2)} - \mathbf{q}_{j-1}^* \mathbf{a}_j \mathbf{q}_{j-1}, & r_{j-1,j} &= \mathbf{q}_{j-1}^* \mathbf{a}_j, \\ \mathbf{q}_j &= \mathbf{q}_j^{(j-1)} / \|\mathbf{q}_j^{(j-1)}\|_2, & r_{jj} &= \|\mathbf{q}_j^{(j-1)}\|_2.\end{aligned}$$

2.2. Modified Gram–Schmidt orthogonalization (MGS)

- MGS is based on the use of

$$\begin{aligned}\tilde{\mathbf{q}}_j &= \mathbf{P}_j \mathbf{a}_j \\ &= (\mathbf{I} - \mathbf{q}_{j-1} \mathbf{q}_{j-1}^*) \cdots (\mathbf{I} - \mathbf{q}_2 \mathbf{q}_2^*) (\mathbf{I} - \mathbf{q}_1 \mathbf{q}_1^*) \mathbf{a}_j\end{aligned}$$

and calculates \mathbf{q}_j by evaluating the following formulas in order:

$$\begin{aligned} \mathbf{q}_j^{(0)} &= \mathbf{a}_j, \\ \mathbf{q}_j^{(1)} &= \mathbf{q}_j^{(0)} - \mathbf{q}_1^* \mathbf{q}_j^{(0)} \mathbf{q}_1, & r_{1j} &= \mathbf{q}_1^* \mathbf{q}_j^{(0)}, \\ \mathbf{q}_j^{(2)} &= \mathbf{q}_j^{(1)} - \mathbf{q}_2^* \mathbf{q}_j^{(1)} \mathbf{q}_2, & r_{2j} &= \mathbf{q}_2^* \mathbf{q}_j^{(1)}, \\ &\vdots & &\vdots \\ \mathbf{q}_j^{(j-1)} &= \mathbf{q}_j^{(j-2)} - \mathbf{q}_{j-1}^* \mathbf{q}_j^{(j-2)} \mathbf{q}_{j-1}, & r_{j-1,j} &= \mathbf{q}_{j-1}^* \mathbf{q}_j^{(j-2)}, \\ \mathbf{q}_j &= \mathbf{q}_j^{(j-1)} / \|\mathbf{q}_j^{(j-1)}\|_2, & r_{jj} &= \|\mathbf{q}_j^{(j-1)}\|_2. \end{aligned}$$

2.3. CGS and MGS algorithms

Algorithm: GS for n linearly independent vectors $\{\mathbf{a}_i\}_{i=1}^n$.

```
for  $j = 1$  to  $n$ 
     $\mathbf{q}_j = \mathbf{a}_j$ 
    for  $i = 1$  to  $j - 1$ 
         $\begin{cases} r_{ij} = \mathbf{q}_i^* \mathbf{a}_j & \text{CGS} \\ r_{ij} = \mathbf{q}_i^* \mathbf{q}_j & \text{MGS} \end{cases}$ 
         $\mathbf{q}_j = \mathbf{q}_j - r_{ij} \mathbf{q}_i$ 
    end
     $r_{jj} = \|\mathbf{q}_j\|_2$ 
     $\mathbf{q}_j = \mathbf{q}_j / r_{jj}$ 
end
```

- The computational cost: $\sim 2mn^2$ (leading term) for $\mathbf{a}_i \in \mathbb{C}^m$
- CGS and MGS are mathematically equivalent. In finite precision arithmetic, MGS introduces smaller errors than CGS.
- Basic Linear Algebra Subroutine (BLAS). CGS is more efficient.

3. QR factorization

- **Definition:** Let m and n be arbitrary positive integers ($m \geq n$ or $m < n$). Given $\mathbf{A} \in \mathbb{C}^{m \times n}$, not necessarily of full rank, a *full QR factorization* of \mathbf{A} is a factorization

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

where $\mathbf{Q} \in \mathbb{C}^{m \times m}$ is unitary, and $\mathbf{R} \in \mathbb{C}^{m \times n}$ is upper triangular. For $m \geq n$, a *reduced QR factorization* of \mathbf{A} is a factorization

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

where $\mathbf{Q}_n \in \mathbb{C}^{m \times n}$ has orthonormal columns, and

$$\mathbf{R}_n = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \ddots & \vdots \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix} \cdot \begin{array}{c} \text{[gray rectangle]} \\ \mathbf{A} \end{array} = \begin{array}{c} \text{[gray rectangle]} \\ \mathbf{Q} \end{array} \cdot \begin{array}{c} \text{[upper triangular rectangle]} \\ \mathbf{R} \end{array}$$

Theorem 9 (Existence of QR)

Every matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ ($m \geq n$) has a reduced QR factorization and a full QR factorization.

Proof.

- Existence of reduced QR factorization.

For the full column rank case, Gram–Schmidt orthogonalization produces a sequence of **reduced** QR factorizations for $\mathbf{A} \in \mathbb{C}^{m \times n}$:

$$\mathbf{A}_j := [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_j] = \mathbf{Q}_j \mathbf{R}_j, \quad j = 1:n.$$

For the rank-deficient case, $\tilde{\mathbf{q}}_j = \mathbf{0}$ at one or more steps j , GS fails to produce \mathbf{q}_j . At this moment, we pick \mathbf{q}_j arbitrarily to be any unit vector orthogonal to $\text{span}\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{j-1}\}$, set $r_{jj} = 0$, and then continue the Gram–Schmidt orthogonalization until we obtain a reduced QR factorization.

- Existence of full QR factorization.

Let $\mathbf{A} = \mathbf{Q}_n \mathbf{R}_n$ be a reduced QR factorization of \mathbf{A} . A full QR factorization can be constructed via

$$\mathbf{A} = \mathbf{QR} := [\mathbf{Q}_n \quad \mathbf{Q}_c] \begin{bmatrix} \mathbf{R}_n \\ \mathbf{0} \end{bmatrix},$$

where $\mathbf{Q}_c \in \mathbb{C}^{m \times (m-n)}$ has orthonormal columns orthogonal to $\text{span}\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$. □

Theorem 10

Every matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ ($m \geq n$) of full column rank has a unique reduced QR factorization $\mathbf{A} = \mathbf{Q}_n \mathbf{R}_n$ with $r_{jj} > 0$.

Proof.

$r_{11}\mathbf{q}_1 = \mathbf{a}_1$ and $r_{11} > 0 \Rightarrow r_{11}$ and \mathbf{q}_1 unique $\Rightarrow r_{12}$ and $r_{22}\mathbf{q}_2$ unique, by $r_{22} > 0 \Rightarrow r_{22}$ and \mathbf{q}_2 unique, and so on. □

3.1. When vectors become continuous functions

- Replace \mathbb{C}^m by $C[-1, 1]$, a linear space of real-valued continuous functions on $[-1, 1]$ with the L^2 inner product

$$\forall f(x), g(x) \in C[-1, 1], \quad \langle f(x), g(x) \rangle_{L^2} = \int_{-1}^1 f(x)g(x)dx,$$

and the norm

$$\|f(x)\|_{L^2} = \sqrt{\langle f(x), f(x) \rangle_{L^2}}.$$

Gram–Schmidt orthogonalization (GS) with respect to the L^2 inner product $\langle f(x), g(x) \rangle_{L^2}$ is: At step j ,

$$\begin{aligned}\tilde{q}_j(x) &= a_j(x) - \sum_{i=1}^{j-1} \langle a_j(x), q_i(x) \rangle_{L^2} q_i(x), \\ q_j(x) &= \tilde{q}_j(x) / \|\tilde{q}_j(x)\|_{L^2}.\end{aligned}$$

The functions $q_j(x)$ satisfy

$$\langle q_i(x), q_j(x) \rangle_{L^2} = \int_{-1}^1 q_i(x) q_j(x) dx = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Then we have “continuous QR factorization”

$$A = QR = \left[\begin{array}{c|c|c|c} & & & \\ q_1(x) & q_2(x) & \cdots & q_n(x) \end{array} \right] \left[\begin{array}{cccc} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & & \vdots \\ & & \ddots & \\ & & & r_{nn} \end{array} \right]$$

where

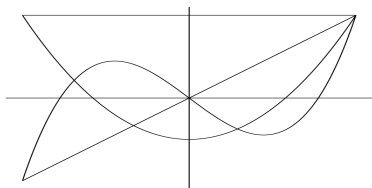
$$A = [a_1(x) \quad a_2(x) \quad \cdots \quad a_n(x)]$$

and

$$r_{jj} = \|\tilde{q}_j(x)\|_{L^2}, \quad r_{ij} = \langle a_j(x), q_i(x) \rangle_{L^2}.$$

- **Example:** $a_j(x) = x^{j-1}$, $j = 1, 2, \dots, n$

$$\mathbf{A} = \begin{bmatrix} 1 & x & x^2 & \dots & x^{n-1} \end{bmatrix}$$



Legendre polynomials $P_j(x) = q_j(x)/q_j(1)$:

$$P_1(x) = x, \quad P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}, \quad P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x.$$

Experiment: Discrete Legendre polynomials

```
x = (-128:128)'/128;
A = [x.^0 x.^1 x.^2 x.^3]
[Q,R] = qr(A,0);
scale = Q(257,:);
Q = Q*diag(1./scale);
plot(x,Q)
```

