# Chap5_Week8b

2022年3月31日 星期四 17:12

- Chap 5.5 Orthogonal Matrix and its properties
- Chap 5.6 Gram-Schmidt Process

A Motivating Example:

Find the orthogonal projection of $\begin{pmatrix}9\\0\\0\\0\end{pmatrix}$ onto span$\left\{\frac{1}{3}\begin{pmatrix}2\\2\\1\\0\end{pmatrix}, \frac{1}{3}\begin{pmatrix}-2\\2\\0\\1\end{pmatrix}\right\}$ $\perp$

$$proj_W\begin{pmatrix}9\\0\\0\\0\end{pmatrix} = \left\langle\begin{pmatrix}9\\0\\0\\0\end{pmatrix}, \frac{1}{3}\begin{pmatrix}2\\2\\1\\0\end{pmatrix}\right\rangle \cdot \frac{1}{3}\begin{pmatrix}2\\2\\1\\0\end{pmatrix} + \left\langle\begin{pmatrix}9\\0\\0\\0\end{pmatrix}, \frac{1}{3}\begin{pmatrix}-2\\2\\0\\1\end{pmatrix}\right\rangle \cdot \frac{1}{3}\begin{pmatrix}-2\\2\\0\\1\end{pmatrix}$$
$$= \frac{1}{9}\cdot 18\begin{pmatrix}2\\2\\1\\0\end{pmatrix} + \frac{1}{9}(-18)\begin{pmatrix}-2\\2\\0\\1\end{pmatrix} = \begin{pmatrix}8\\0\\2\\-2\end{pmatrix}$$

## Def 5.5.12 Orthogonal Matrix

An $n\times n$ matrix $Q$ is said to be orthogonal if the columns in $Q$ form an orthonormal set in $\mathbb{R}^n$.

→ also an orthonormal BASIS of $\mathbb{R}^n$

Example

Rotation matrix

$$Q = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

is orthogonal and

$$Q^{-1} = Q^T = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}.$$

(Rotation preserves length of vector and angle between 2 vectors (Example 4.1.16).)

$\|Q\vec{x}\| = \|\vec{x}\|$

$$\frac{\langle Q\vec{x}, Q\vec{y}\rangle}{\|Q\vec{x}\|\,\|Q\vec{y}\|} = \frac{\langle\vec{x},\vec{y}\rangle}{\|\vec{x}\|\,\|\vec{y}\|}$$

Example

Permutation matrix is a matrix formed from the identity by reordering its column.

$$\begin{pmatrix}1&&\\&1&\\&&1\end{pmatrix},\begin{pmatrix}1&&\\&&1\\&1&\end{pmatrix},\begin{pmatrix}&1&\\1&&\\&&1\end{pmatrix},\begin{pmatrix}&1&\\&&1\\1&&\end{pmatrix},\begin{pmatrix}&&1\\1&&\\&1&\end{pmatrix},\begin{pmatrix}&&1\\&1&\\1&&\end{pmatrix}$$

3×3 permutation matrices

Permutation matrices are orthogonal.

**Thm 5.5.13 (3-4)**

$$\begin{cases}\langle Q\vec{x}, Q\vec{y}\rangle = \langle\vec{x},\vec{y}\rangle\\ \|Q\vec{x}\| = \|\vec{x}\|\end{cases}$$

**Thm 5.5.13 (1-2)**

$Q$ is an orthogonal matrix $\Leftrightarrow$ $Q^{-1} = Q^T$

$$\underset{\text{orthogonal}}{Q} = \begin{bmatrix} | & | & & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ | & | & & | \end{bmatrix}_{n\times n} \quad \text{then } Q^T = \begin{bmatrix} — & \vec{q}_1^{\,T} & — \\ — & \vec{q}_2^{\,T} & — \\ & \vdots & \\ — & \vec{q}_n^{\,T} & — \end{bmatrix}_{n\times n}$$

$Q^T_{n\times n} Q_{n\times n} = M_{n\times n}$, with $M_{ij} = \langle\vec{q}_i, \vec{q}_j\rangle = \begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}$

---

Chap 5.6

1. Master the skill of using Gram-Schmidt Process to get an orthonormal basis
2. Know QR factorization as an application to the G-S process.

Input : A basis $\{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p\}$

↓ G-S ?

Output : An orthonormal basis $\{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_p\}$

An orthogonal basis $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_p\}$

Section 5.6 Gram-Schmidt Process

Theorem 5.6.1 (Gram-Schmidt Process)

Given a basis $\{x_1, x_2, \cdots, x_n\}$ for a nonzero subspace $W$ of $V$, define

Stage I: Orthogonalization

$$v_1 = x_1$$
$$v_2 = x_2 - \frac{\langle x_2, v_1\rangle}{\langle v_1, v_1\rangle}v_1$$
$$v_3 = x_3 - \frac{\langle x_3, v_1\rangle}{\langle v_1, v_1\rangle}v_1 - \frac{\langle x_3, v_2\rangle}{\langle v_2, v_2\rangle}v_2$$
$$\vdots$$
$$v_n = x_n - \frac{\langle x_n, v_1\rangle}{\langle v_1, v_1\rangle}v_1 - \frac{\langle x_n, v_2\rangle}{\langle v_2, v_2\rangle}v_2 - \cdots - \frac{\langle x_n, v_{n-1}\rangle}{\langle v_{n-1}, v_{n-1}\rangle}v_{n-1}$$

Stage II: Normalization

Then $\left\{\frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \cdots, \frac{v_n}{\|v_n\|}\right\}$ is an orthonormal basis for $W$. In addition

$$span\{v_1, v_2, \cdots, v_k\} = span\{x_1, x_2, \cdots, x_k\} \quad \text{for } 1\le k\le n.$$

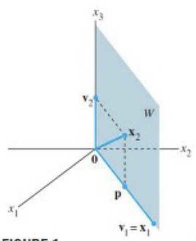



**FIGURE 1** Construction of an orthogonal basis $\{v_1, v_2\}$.

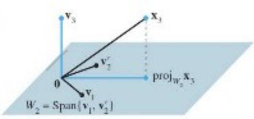



**FIGURE 2** The construction of $v_3$ from $x_3$ and $W_2$.

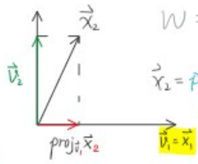
## Gram-Schmidt Process :

Given $\{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n\}$ as a basis,

Stage I: Orthogonalization

Step 1 : $\vec{v}_1 = \vec{x}_1$

Step 2 : $\vec{v}_2 =$




$W = span\{\vec{v}_1, \vec{v}_2\}$

$\vec{x}_2 = proj_{\vec{v}_1}\vec{x}_2 + \vec{v}_2$

$\in span\{\vec{v}_1\}$ ; $\in span\{\vec{v}_1\}^\perp$

$\vec{v}_2 \perp \vec{v}_1$

$\Rightarrow$ $\vec{v}_2 \perp \vec{v}_1$ and $span\{\vec{v}_1, \vec{v}_2\} = span\{\vec{x}_1, \vec{x}_2\} = W$

Step 3 : $\vec{v}_3 = \vec{x}_3 - proj_W\vec{x}_3$

$$= \vec{x}_3 - proj_{\vec{v}_1}\vec{x}_3 - proj_{\vec{v}_2}\vec{x}_3 = \vec{x}_3 - \frac{\langle\vec{x}_3, \vec{v}_1\rangle}{\langle\vec{v}_1, \vec{v}_1\rangle}\vec{v}_1 - \frac{\langle\vec{x}_3, \vec{v}_2\rangle}{\langle\vec{v}_2, \vec{v}_2\rangle}\vec{v}_2$$

$\Rightarrow$ $\vec{v}_3 \perp \vec{v}_1$, $\vec{v}_3 \perp \vec{v}_2$ and $span\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = span\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$

...

↳ An orthogonal basis $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_p\}$

Stage II : Normalization

$\vec{u}_i = \vec{v}_i / \|\vec{v}_i\|$ , $i = 1, 2, \cdots, p$.

↳ An orthonormal basis $\{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_p\}$.

**E.g.** Let $W = Span\left\{\underset{\vec{x}_1}{\begin{pmatrix}1\\0\\1\end{pmatrix}}, \underset{\vec{x}_2}{\begin{pmatrix}1\\1\\1\end{pmatrix}}\right\}$ be a subspace of $\mathbb{R}^3$

Use G-S process to obtain an orthonormal basis of $W$.

Stage I : Take $\vec{v}_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}$, then

$$\vec{v}_2 = \begin{pmatrix}1\\1\\1\end{pmatrix} - proj_{\vec{v}_1}\vec{x}_2 = \begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{(1\ 1\ 1)\begin{pmatrix}1\\0\\1\end{pmatrix}}{(1\ 0\ 1)\begin{pmatrix}1\\0\\1\end{pmatrix}}\begin{pmatrix}1\\0\\1\end{pmatrix}$$
$$= \begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\end{pmatrix}$$

Stage II : $\vec{u}_1 = \vec{v}_1/\|\vec{v}_1\| = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}$ , $\vec{u}_2 = \vec{v}_2/\|\vec{v}_2\| = \begin{pmatrix}0\\1\\0\end{pmatrix}$.

$W = span\left\{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right\}$.

$$W = \text{span}\left\{ \tfrac{1}{\sqrt2}\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix} \right\}.$$

**Exercise:** The same problem as before. Start with $\vec v_1 = \vec x_2$.

Stage I: Take $\vec v_1 = \vec x_2 = \begin{pmatrix}1\\1\\1\end{pmatrix}$,

$$\vec v_2 = \vec x_1 - \text{proj}_{\vec u_1}\vec x_1 = \begin{pmatrix}1\\0\\1\end{pmatrix} - \frac{2}{3}\begin{pmatrix}1\\1\\1\end{pmatrix} = \frac{1}{3}\begin{pmatrix}1\\-2\\1\end{pmatrix}$$

Stage II: $\{\vec u_1, \vec u_2\} = \left\{ \frac{1}{\sqrt3}\begin{pmatrix}1\\1\\1\end{pmatrix}, \frac{\sqrt6}{3}\begin{pmatrix}1\\-2\\1\end{pmatrix} \right\}$

**Observations:**

1. What happens if $\vec v_1$ is chosen differently? Would $\{\vec u_1, \cdots, \vec u_p\}$ change? Yes
2. What happens if the inner product is defined differently? Would $\{\vec u_1, \cdots, \vec u_p\}$ change? Yes

**Remark:** The default choice of $\vec v_1$ is the first vector in the set.

Example

Let $x_1 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$, $x_2 = \begin{pmatrix}0\\1\\1\\1\end{pmatrix}$ and $x_3 = \begin{pmatrix}0\\0\\1\\1\end{pmatrix}$. Construct an orthonormal basis for a subspace $span\left\{\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\\1\end{pmatrix}\right\}$ of $R^4$.

I: Solution

Step 1: Let $v_1 = x_1$ and $W_1 = span\{v_1\} = span\{x_1\}$.

2: Let $v_2 = x_2 - \text{proj}_{W_1} x_2 = x_2 - \frac{\langle x_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = \begin{pmatrix}0\\1\\1\\1\end{pmatrix} - \frac{3}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \begin{pmatrix}-3/4\\1/4\\1/4\\1/4\end{pmatrix}$.

$\{v_1, v_2\}$ is an orthogonal basis for the subspace $W_2 = \{x_1, x_2\}$.

$$\text{proj}_{W_2} x_3 = \underbrace{\frac{\langle x_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1}_{\substack{\text{projection of}\\ x_3 \text{ onto } v_1}} + \underbrace{\frac{\langle x_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2}_{\substack{\text{projection of}\\ x_3 \text{ onto } v_2}} = \frac{2}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + \frac{\frac{2}{4}}{\frac{12}{16}}\begin{pmatrix}-3/4\\1/4\\1/4\\1/4\end{pmatrix} = \begin{pmatrix}0\\2/3\\2/3\\2/3\end{pmatrix}.$$

3: Let $v_3 = x_3 - \text{proj}_{W_2} x_3 = \begin{pmatrix}0\\0\\1\\1\end{pmatrix} - \begin{pmatrix}0\\2/3\\2/3\\2/3\end{pmatrix} = \begin{pmatrix}0\\-2/3\\1/3\\1/3\end{pmatrix}$.

$\{v_1, v_2, v_3\}$ is an orthogonal basis for $W$.

II: Normalizing $\{v_1, v_2, v_3\}$, $\left\{\frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \frac{1}{\sqrt{12}}\begin{pmatrix}-3\\1\\1\\1\end{pmatrix}, \frac{1}{\sqrt6}\begin{pmatrix}0\\-2\\1\\1\end{pmatrix}\right\}$ is an orthonormal basis for $W$.

---

**Exercise:** Find the orthogonal projection of $\begin{pmatrix}9\\0\\0\\0\end{pmatrix} = \vec b$ onto $W = \text{span}\left\{\begin{pmatrix}0\\4\\1\\1\end{pmatrix}, \begin{pmatrix}1\\3\\0\\1/2\end{pmatrix}\right\}$.

idea 1: Use G-S process on $\left\{\begin{pmatrix}0\\4\\1\\1\end{pmatrix}, \begin{pmatrix}1\\3\\0\\1/2\end{pmatrix}\right\}$, then find $\text{proj}_W\begin{pmatrix}9\\0\\0\\0\end{pmatrix}$

G-S: Stage I: Taking $\vec v_1 = \begin{pmatrix}0\\4\\1\\1\end{pmatrix}$, then

$$\vec v_2 = \begin{pmatrix}1\\3\\0\\1/2\end{pmatrix} - \text{proj}_{\vec v_1}\begin{pmatrix}1\\3\\0\\1/2\end{pmatrix} = \begin{pmatrix}1\\3\\0\\1/2\end{pmatrix} - \frac{25/2}{18}\begin{pmatrix}0\\4\\1\\1\end{pmatrix} = \begin{pmatrix}1\\3\\0\\1/2\end{pmatrix} - \frac{25}{36}\begin{pmatrix}0\\4\\1\\1\end{pmatrix} = ?$$

Stage II: $\vec u_1 = \vec v_1 / \|\vec v_1\| = \frac{1}{\sqrt{18}}\begin{pmatrix}0\\4\\1\\1\end{pmatrix}$, $\vec u_2 = \vec v_2 / \|\vec v_2\| = ?$

Projection: $\text{proj}_W\begin{pmatrix}9\\0\\0\\0\end{pmatrix} = \left\langle \begin{pmatrix}9\\0\\0\\0\end{pmatrix}, \vec u_1 \right\rangle \vec u_1 + \left\langle \begin{pmatrix}9\\0\\0\\0\end{pmatrix}, \vec u_2 \right\rangle \vec u_2 = ?$

idea 2: Let $A = \begin{pmatrix}0 & 1\\4 & 3\\1 & 0\\1 & 1/2\end{pmatrix}$, then $A(A^TA)^{-1}A^T\vec b$

---

## △ QR - factorization ** (optional)

Theorem 5.6.1 ($QR$ factorization)

Use the same notation as in Theorem 5.6.1 where $V = R^m$. If $A = (\boldsymbol{x}_1 \mid \boldsymbol{x}_2 \mid \cdots \mid \boldsymbol{x}_n)$ is an $m \times n$ matrix of rank $n$, then $A$ can be factored into a product $QR$, where

$$Q = \left( \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|} \middle| \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|} \middle| \cdots \middle| \frac{\boldsymbol{v}_n}{\|\boldsymbol{v}_n\|} \right) \text{ and } R = \begin{pmatrix} \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}\cdot \boldsymbol{v}_1 & \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}\cdot \boldsymbol{x}_2 & \cdots & \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}\cdot \boldsymbol{x}_n \\ 0 & \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|}\cdot \boldsymbol{v}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{\boldsymbol{v}_{n-1}}{\|\boldsymbol{v}_{n-1}\|}\cdot \boldsymbol{x}_n \\ 0 & \cdots & 0 & \frac{\boldsymbol{v}_n}{\|\boldsymbol{v}_n\|}\cdot \boldsymbol{v}_n \end{pmatrix}$$

QR factorization is often used in improving computational efficiency and fast computation with computers. It saves some computational efforts, storage and reduces the time cost.

You will see how G-S process being applied in QR factorization here. But since the computations in QR factorization will be quite tedious and we omitted the proof and you may read it as a complementary material.

Example 5.6.2

Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ and $Q = \begin{pmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{pmatrix}$ as in Example 5.5.2. We have

$$\frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}\cdot \boldsymbol{v}_1 = 2 \quad \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}\cdot \boldsymbol{x}_2 = 3/2 \quad \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}\cdot \boldsymbol{x}_3 = 1$$

$$\frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|}\cdot \boldsymbol{v}_2 = \sqrt{3}/2 \quad \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|}\cdot \boldsymbol{x}_3 = 1/\sqrt{3}$$

$$\frac{\boldsymbol{v}_3}{\|\boldsymbol{v}_3\|}\cdot \boldsymbol{v}_3 = \sqrt{2}/\sqrt{3}$$

Then $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 2 & 3/2 & 1 \\ 0 & \sqrt{3}/2 & 1/\sqrt{3} \\ 0 & 0 & \sqrt{2}/\sqrt{3} \end{pmatrix}$.

Theorem 5.6.3

If $A$ is an $m \times n$ matrix of rank $n$, then the least squares solution of $A\boldsymbol{x} = \boldsymbol{b}$ is given by $\hat{\boldsymbol{x}} = R^{-1}Q^T\boldsymbol{b}$, where $Q$ and $R$ are the matrices obtained from $QR$ factorization.

---

Gram-Schmidt in Matlab (you may copy the code after'>>')

Code for example 2:

```
>>  x1=ones(4,1); x2=[0;1;1;1]; x3=[0;0;1;1];
>>  v1=x1  % take v1=x1
>>  v2=x2-(x2'*v1)/(v1'*v1)*v1   % obtain v2
>>  v3=x3-(x3'*v1)/(v1'*v1)*v1-(x3'*v2)/(v2'*v2)*v2   %obtain v3
>>  u1=v1/norm(v1); u2=v2/norm(v2); u3=v3/norm(v3);  %Normalize
```

Modify the program and apply on another example

Eg. $W = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$