Lecture 24: Quick review from previous lecture

- A square matrix A is called an **orthogonal matrix** if it satisfies

$$A^T A = AA^T = I. \quad \Rightarrow \quad A^{-1} = A^T$$

- A is an **orthogonal matrix** if and only if its columns (rows) form an orthonormal basis on $\mathbb{R}^n$ with respect to the Euclidean dot product.
- If A and B are **orthogonal matrices**, then AB is **orthogonal** too.

Today we will discuss

- Sec. 4.3 Orthogonal Matrices and QR factorization
- Sec. 4.4 Orthogonal Projections

- Lecture will be recorded -

---

§ The QR Factorization

Q: How do we turn the matrix $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ to an orthogonal matrix?

Answer: This can be achieved by applying the Gram-Schmidt equation.

Let's revisit the Gram-Schmidt process:

Let

$$A = [\mathbf{a}_1 \cdots \mathbf{a}_n] \text{ be } n \times n \text{ nonsingular matrix}$$

where $\mathbf{a}_j$ is the i^{th} column vector of A. Thus, $\mathbf{a}_1, \cdots, \mathbf{a}_n$ are linearly independent.

Step 1: Use Gram-Schmidt to turn linearly independent $\mathbf{a}_1, \cdots, \mathbf{a}_n$ to orthogonal vectors $\mathbf{v}_1, \cdots, \mathbf{v}_n$.

Step 2: Normalize $\mathbf{v}_1, \cdots, \mathbf{v}_n$ to get orthonormal vectors $\mathbf{q}_1, \cdots, \mathbf{q}_n$:
That is,

$$\mathbf{q}_j = \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}$$

Then a nonsingular matrix A is turned to an orthogonal matrix Q:

$$\underbrace{\left(\begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array} \right)}_{A \text{ (nonsingular)}} \underbrace{\xrightarrow{\text{Step 1}}}_{GramSchmidt} \underbrace{\left(\begin{array}{c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{array} \right)}_{V}$$

$$\underbrace{\xrightarrow{\text{Step 2}}}_{Normalization} \underbrace{\left(\begin{array}{c|c|c|c} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{array} \right)}_{Q \text{ (orthogonal)}}$$

To answer the earlier question:

Example 3. Now let's turn the matrix $A = \begin{array}{cc} a_1 & a_2 \\ \hline \end{array} \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ to an orthogonal matrix.

1. By G-S process,

$$\begin{cases} v_1 = a_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ v_2 = a_2 - \dfrac{\langle a_2, v_1 \rangle}{\|v_1\|^2} v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \dfrac{\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \rangle}{\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \rangle} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{cases}$$

$$= \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{3}{5} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2/5 \\ -1/5 \end{pmatrix}$$

$\rightarrow \|v_2\| = \sqrt{(\frac{2}{5})^2 + (\frac{1}{5})^2}$

$\begin{pmatrix} 2 \\ -1 \end{pmatrix}$

2. $q_1 = \dfrac{v_1}{\|v_1\|} = \dfrac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$; $q_2 = \dfrac{v_2}{\|v_2\|} = \dfrac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix}$

$A \longrightarrow Q = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}$ (orthogonal matrix)

Q: From above example, we have seen that $A \to Q$. Indeed, we will be able to decompose

$$A = Q\blacksquare.$$

What is this matrix $\blacksquare$ that encodes all processes turning A to Q? This matrix $\blacksquare$ is actually upper triangular.

$c\langle x, y \rangle = \langle x, cy \rangle$

Let's figure this out.

In Step 1 above, we did Gram-Schmidt process:

$q_1 = \dfrac{v_1}{\|v_1\|}$

$$r_{11} q_1 = \mathbf{v}_1 = \mathbf{a}_1$$

$$r_{22} q_2 \quad \mathbf{v}_2 = \mathbf{a}_2 - \boxed{\frac{\langle \mathbf{a}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1} = \langle a_2, \underbrace{\frac{v_1}{\|v_1\|}}_{= q_1} \rangle \underbrace{\frac{v_1}{\|v_1\|}}_{= q_1}$$

$$\mathbf{v}_3 = \mathbf{a}_3 - \frac{\langle \mathbf{a}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\langle \mathbf{a}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2$$

$$\vdots$$

$$\mathbf{v}_n = \mathbf{a}_n - \frac{\langle \mathbf{a}_n, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \cdots - \frac{\langle \mathbf{a}_n, \mathbf{v}_{n-1} \rangle}{\|\mathbf{v}_{n-1}\|^2} \mathbf{v}_{n-1}$$

$$\boxed{\mathbf{q}_j = \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|} \underbrace{\Rightarrow}_{r_{jj}=\|\mathbf{v}_j\|} r_{jj}\mathbf{q}_j = \mathbf{v}_j}$$

We can get

$$\begin{aligned}
r_{11}\mathbf{q}_1 &= \mathbf{a}_1 \\
r_{22}\mathbf{q}_2 &= \mathbf{a}_2 - \overbrace{\langle \mathbf{a}_2, \mathbf{q}_1 \rangle}^{= r_{12}} \mathbf{q}_1 \\
r_{33}\mathbf{q}_3 &= \mathbf{a}_3 - \underbrace{\langle \mathbf{a}_3, \mathbf{q}_1 \rangle}_{r_{13}} \mathbf{q}_1 - \underbrace{\langle \mathbf{a}_3, \mathbf{q}_2 \rangle}_{r_{23}} \mathbf{q}_2 \\
&\vdots \\
r_{nn}\mathbf{q}_n &= \mathbf{a}_n - \langle \mathbf{a}_n, \mathbf{q}_1 \rangle \mathbf{q}_1 - \cdots - \langle \mathbf{a}_n, \mathbf{q}_{n-1} \rangle \mathbf{q}_{n-1}
\end{aligned}$$

Let $r_{ij} = \langle \mathbf{a}_j, \mathbf{q}_i \rangle$, then we reach

$$\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\
\mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\
\mathbf{a}_3 &= r_{13}\mathbf{q}_1 + r_{23}\mathbf{q}_2 + r_{33}\mathbf{q}_3 \\
&\vdots \\
\mathbf{a}_n &= r_{1n}\mathbf{q}_1 + +r_{2n}\mathbf{q}_2 + \cdots + r_{n-1,n}\mathbf{q}_{n-1} + r_{nn}\mathbf{q}_n
\end{aligned}$$

Thus the Gram-Schmidt equations can be recast into an equivalent matrix form:

$$\underbrace{\left(\begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array}\right)}_{A \text{ (nonsingular)}} = \underbrace{\left(\begin{array}{c|c|c|c} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{array}\right)}_{Q \text{ (orthogonal)}} \underbrace{\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{pmatrix}}_{R}$$

where

$$r_{jj} = \|\mathbf{v}_j\| = \langle \mathbf{a}_j, \mathbf{q}_j \rangle \qquad \text{and} \qquad r_{ij} = \langle \mathbf{a}_j, \mathbf{q}_i \rangle.$$

This is called the **QR factorization**.

"matrix form" of G-S process.

In conclusion, we have

Fact 6: Every nonsingular matrix can be factored into

$$\boxed{A = QR}$$

the product of an **orthogonal** matrix Q and an **upper triangular** matrix R.

Back to **Example 3:** Find QR factorization of $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ under the usual dot product.

$$R = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix}, = \begin{bmatrix} \sqrt{5} & 3/\sqrt{5} \\ 0 & 1/\sqrt{5} \end{bmatrix}$$

$$r_{11} = \|v_1\| = \sqrt{5}.$$

$$r_{12} = \langle a_2, q_1 \rangle = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \right\rangle = 3/\sqrt{5}.$$

$$r_{22} = \|v_2\| = 1/\sqrt{5}.$$

$$\text{Then } A = QR \quad \#$$

$$v_2 = \begin{pmatrix} 2/5 \\ -1/5 \end{pmatrix}, \quad \|v_2\| = \sqrt{\frac{4+1}{25}} = \sqrt{\frac{1}{5}}$$

Example 4. Let $A = [\mathbf{w}_1\ \mathbf{w}_2\ \mathbf{w}_3]$ where

$$\mathbf{w}_1 = (1, 1, 0)^T, \quad \mathbf{w}_2 = (0, 1, 1)^T, \quad \mathbf{w}_3 = (1, 0, 1)^T.$$

Find the QR factorization of A under the usual dot product.

[**Answer:**] **(1) Find** Q**.** We have seen in **Lecture 22, Example 1** how to apply Gram-Schmidt to turn vectors $\mathbf{w}_i$ into an orthonormal basis:

$$\mathbf{q}_1 = \frac{1}{\sqrt{2}} \underbrace{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}_{\mathbf{v}_1}, \mathbf{q}_2 = \sqrt{\frac{2}{3}} \underbrace{\begin{pmatrix} -1/2 \\ 1/2 \\ 1 \end{pmatrix}}_{\mathbf{v}_2}, \mathbf{q}_3 = \frac{\sqrt{3}}{2} \underbrace{\begin{pmatrix} 2/3 \\ -2/3 \\ 2/3 \end{pmatrix}}_{\mathbf{v}_3}.$$

$$Q = [q_1\ q_2\ q_3]$$

Specifically,

$$\underbrace{\left(\begin{array}{c|c|c} \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \end{array}\right)}_{A} \underbrace{\longrightarrow}_{GramSchmidt} \underbrace{\left(\begin{array}{c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{array}\right)}_{V} \underbrace{\longrightarrow}_{Normalization} \underbrace{\left(\begin{array}{c|c|c} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \end{array}\right)}_{Q \text{ (orthogonal)}}$$

(2) It is remaining to find R. That is, $R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix}$ with $r_{jj} = \|\mathbf{v}_j\| = \langle \mathbf{a}_j, \mathbf{q}_j \rangle$ and $r_{ij} = \langle \mathbf{a}_j, \mathbf{q}_i \rangle$.

$r_{11} = \| v_1 \| = \left\| \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\| = \sqrt{2}$,

$r_{22} = \| v_2 \| = \left\| \begin{pmatrix} -1/2 \\ 1/2 \\ 1 \end{pmatrix} \right\| = \sqrt{3/2}$.

$r_{33} = \| v_3 \| = 2/\sqrt{3}$.

$r_{12} = \langle w_2, q_1 \rangle = \left\langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \right\rangle = 1/\sqrt{2}$.

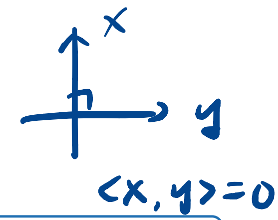
$r_{13} = \langle w_3, q_1 \rangle = 1/\sqrt{2}$.

$r_{23} = \langle w_3, q_2 \rangle = 1/\sqrt{6}$.

Then $A = QR$, where

$$R = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3/2} & 1/\sqrt{6} \\ 0 & 0 & 2/\sqrt{3} \end{bmatrix}. \quad \#$$

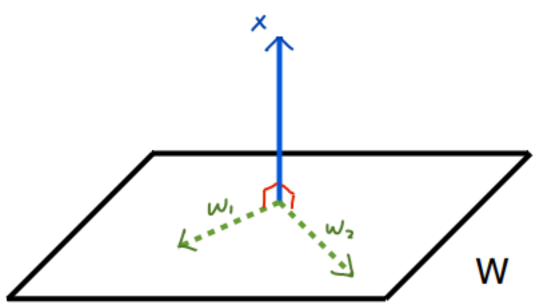
4.4 Orthogonal Projections and Subspaces



§ Orthogonal projections

Definition: We call a vector $\mathbf{x}$ in an inner product space V is **orthogonal** to the subspace W of V if it is orthogonal to every vector in W, that is,

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0 \qquad \text{for all } \mathbf{y} \in W.$$



$\langle x, w_1 \rangle = 0$
$\langle x, w_2 \rangle = 0$

Fact 1: Suppose $\mathbf{w}_1, \cdots, \mathbf{w}_n$ is the basis of W. Thus,

$$\mathbf{x} \text{ is orthogonal to } W \Leftrightarrow \langle \mathbf{x}, \mathbf{w}_i \rangle = 0,\ i = 1, \cdots, n.$$

[To see this:] $(\Rightarrow)$ Since $w_i \in W$, $\langle x, w_i \rangle = 0$, $1 \le i \le n$.

$(\Leftarrow)$ Any $z \in W$, we can write

$z = c_1 w_1 + \cdots + c_n w_n$

$\langle x, z \rangle = \langle x, c_1 w_1 + \cdots + c_n w_n \rangle$

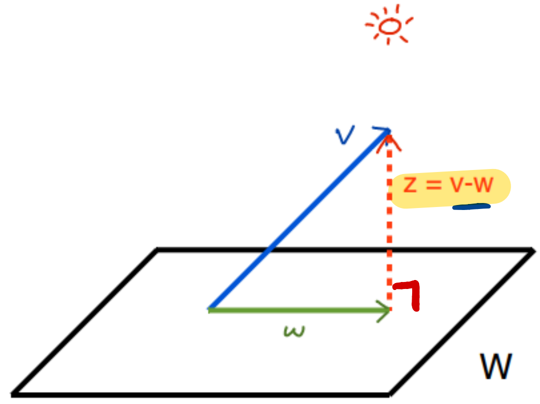
$= c_1 \langle x, w_1 \rangle + \cdots + c_n \langle x, w_n \rangle$

$= 0$ since $\langle x, w_i \rangle = 0$, $1 \le i \le n$.

Definition: The **orthogonal projection** of $\mathbf{v}$ onto the subspace W of V is the element $\mathbf{w} \in W$ such that the difference $\mathbf{z} = \mathbf{v} - \mathbf{w}$ orthogonal to W. We use the notation

$$\mathbf{z} \perp W.$$

The orthogonal projection is "unique". Note that such $\mathbf{w}$ is the unique vector in W that is "closet to" $\mathbf{v}$.



w : orthogonal projection of v onto space W

$$w = \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \cdots + \frac{\langle v, v_n \rangle}{\|v_n\|^2} v_n$$

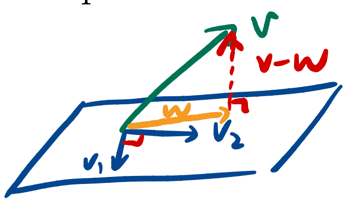
Fact 2: (1) Suppose $\mathbf{v}_1, \cdots, \mathbf{v}_n$ is an orthogonal basis of subspace W of V. If $\mathbf{w} \in W$ is the orthogonal projection of $\mathbf{v} \in V$ onto W, then

$$\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n, \qquad \text{where } c_j = \frac{\langle \mathbf{v}, \mathbf{v}_j \rangle}{\|\mathbf{v}_j\|^2}, \quad j = 1, \cdots, n.$$

(2) Suppose $\mathbf{u}_1, \cdots, \mathbf{u}_n$ is an orthonormal basis of subspace W of V. If $\mathbf{w} \in W$ is the orthogonal projection of $\mathbf{v} \in V$ onto W, then

$$\mathbf{w} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n, \qquad \text{where } c_j = \langle \mathbf{v}, \mathbf{u}_j \rangle, \quad j = 1, \cdots, n.$$

Example 1.[p.214 in Textbook] Suppose that the plane is spanned by the orthogonal vectors $\mathbf{v}_1 = (1, -2, 1)^T$ and $\mathbf{v}_2 = (1, 1, 1)^T$ under the usual dot product. Compute the orthogonal projection of $\mathbf{v} = (1, 0, 0)^T$.



$$w = \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2} v_2 .$$

$$= \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \end{pmatrix} . \#$$

$$v - w = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \\ -1/2 \end{pmatrix}, \quad \langle v-w, v_1 \rangle = \left\langle \begin{pmatrix} 1/2 \\ 0 \\ -1/2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\rangle = 0$$

$$\langle v-w, v_2 \rangle = \left\langle \begin{pmatrix} 1/2 \\ 0 \\ -1/2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle = 0.$$

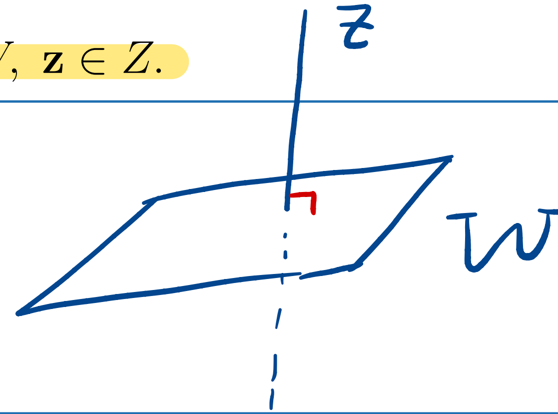
Remark: Thus, $\mathbf{v} - \sum_{k=1}^{n} \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k$ is orthogonal to W, that is,

$$\left(\mathbf{v} - \underbrace{\sum_{k=1}^{n} \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k}_{w}\right) \perp W.$$

§ Orthogonal Subspaces

Definition: Two subspaces W, Z of V are called **orthogonal** if every vector in W is orthogonal to every vector in Z, that is,

$$\langle \mathbf{w}, \mathbf{z} \rangle = 0 \quad \text{for all } \mathbf{w} \in W,\ \mathbf{z} \in Z.$$



Immediately, we also have

Fact 3 : If $\{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$ span W and $\{\mathbf{z}_1, \cdots, \mathbf{z}_k\}$ span Z, then

$$W, Z \text{ are } \textbf{orthogonal} \iff \langle \mathbf{w}_i, \mathbf{z}_j \rangle = 0$$

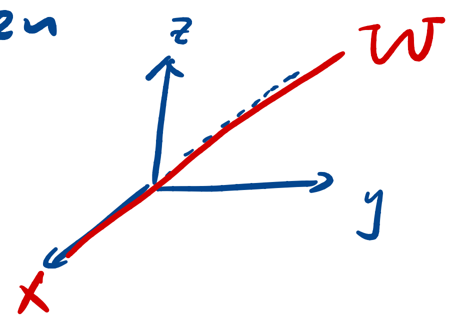
for all $1 \leq i \leq n$, $1 \leq j \leq k$.

Definition: If W is a subspace of V, its **orthogonal complement** $W^\perp$ (pronounced "W perp") is the set of all vectors orthogonal to W, that is,

$$W^\perp = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ for all } \mathbf{w} \in W\}.$$

EX: $W = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$. Then

$W^\perp = $? yz plane



- It can be checked that $W^\perp$ is also a subspace of V.
- If $W = \text{span}\{\mathbf{w}\}$, we will also denote $W^\perp$ by $\mathbf{w}^\perp$.
- Note that the "only vector" contained in both W and $W^\perp$ is $\mathbf{0}$.