Lecture 24: Quick review from previous lecture

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

$$A^T A = AA^T = I.$$

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

Today we will discuss the QR factorization and orthogonal projections.

- Lecture will be recorded -

Lecture video can be found in Canvas "Media Gallery"

§ The QR Factorization

Revisit the Gram-Schmidt process:

Let $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$ be $n \times n$ 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.

Turn $\mathbf{a}_1, \cdots, \mathbf{a}_n$ to orthogonal vectors $\mathbf{v}_1, \cdots, \mathbf{v}_n$:

Gram-Schmidt

$$\begin{cases} \underbrace{\mathbf{v}_1}_{\|v_1\| q_1} = \mathbf{a}_1 \\ \underbrace{\mathbf{v}_2}_{\|v_2\| q_2} = \mathbf{a}_2 - \underbrace{\dfrac{\langle \mathbf{a}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1}_{\langle a_2, q_1 \rangle q_1} \\ \mathbf{v}_3 = \mathbf{a}_3 - \underbrace{\dfrac{\langle \mathbf{a}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1}_{\langle a_3, q_1 \rangle q_1} - \underbrace{\dfrac{\langle \mathbf{a}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2}_{= \langle a_3, q_2 \rangle q_2.} \\ \quad \vdots \\ \mathbf{v}_n = \mathbf{a}_n - \dfrac{\langle \mathbf{a}_n, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \cdots - \dfrac{\langle \mathbf{a}_n, \mathbf{v}_{n-1} \rangle}{\|\mathbf{v}_{n-1}\|^2}\mathbf{v}_{n-1} \end{cases}$$

$q_j = v_j / \|v_j\|$

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

That is,

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

$v_j = \|v_j\| q_j .$

$= r_{jj} \; q_j .$

We can get

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

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

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

We reach $\{a_1, \ldots, a_n\}$ l. independent $\overset{GS}{\Longrightarrow} \{v_1, \ldots, v_n\}$ orthogonal $\overset{q_i = v_i/\|v_i\|}{\Longrightarrow} \{q_1, \ldots, q_n\}$ ON

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

The Gram-Schmidt equations can then be recast into an equivalent matrix form:

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

where $r_{jj} = \|\mathbf{v}_j\|$ and $r_{ij} = \langle \mathbf{a}_j, \mathbf{q}_i \rangle$. This is called the **QR factorization**.

$= \langle a_j, q_j \rangle$

Q is orthogonal

Fact: Every nonsingular matrix can be factored into, $A = QR$, the product of an **orthogonal** matrix Q and an **upper triangular** matrix R.

The factorization is unique if R is positive upper triangular(meaning that all its diagonal entries are positive ($r_{jj} > 0$).

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

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Find the QR factorization of A.

[Ans:] (1) We have saw in Lecture 22 how to apply Gram-Schmidt to the vectors $\mathbf{a}_i$: We found the resulting orthogonal basis to be

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

Thus, letting $\mathbf{q}_k = \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|}$, we have the **orthonormal basis** is

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

It is remaining to find R.

We have known $Q = [\, q_1 \;\; q_2 \;\; q_3 \,]$.

To find R: $R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix}$.

$$\begin{cases} r_{ii} = \|v_i\|. \\ r_{ij} = \langle a_j, q_i \rangle \end{cases}$$

$r_{11} = \|v_1\| = \sqrt{2}$, $r_{22} = \|v_2\| = \sqrt{\frac{3}{2}}$

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

[Example Continue:]

$r_{12} = \langle a_2, q_1 \rangle = \left\langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle = \frac{1}{\sqrt{2}}.$

$r_{13} = \langle a_3, q_1 \rangle = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle = \frac{1}{\sqrt{2}}.$

$r_{23} = \langle a_3, q_2 \rangle = 1/\sqrt{6}.$

$R = \begin{pmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{pmatrix}.$ #

§ Use QR factorization of A to solve linear system $A\mathbf{x} = \mathbf{b}$.

If $A = QR$, we want to use it to solve the system $A\mathbf{x} = \mathbf{b}$:
From the linear system, we get

$$QR\mathbf{x} = \mathbf{b}$$

Since $Q^{-1} = Q^T$ (Q is orthogonal), we have

$$R\mathbf{x} = Q^T\mathbf{b}.$$

Therefore, we just need to solve

$$\boxed{R\mathbf{x} = Q^T\mathbf{b}}$$

by applying back substitution since R is upper triangular matrix.

Example. Apply this to solve the system

$$\overbrace{\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}}^{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \mathbf{b}$$

From previous example, $A = QR$.

$QR \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$

$R \begin{pmatrix} x \\ y \\ z \end{pmatrix} = Q^T \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

[Example Continue:]

$$\begin{pmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{3}{2}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

$$= \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

Using back substitution,

① $\frac{2}{\sqrt{3}} z = \frac{1}{\sqrt{3}}$. $\Rightarrow z = \frac{1}{2}$.

② $\sqrt{\frac{3}{2}} y + \frac{1}{\sqrt{6}} z = \frac{-1}{\sqrt{6}} \Rightarrow y = -1/2$.

③ $x = 1/2$.

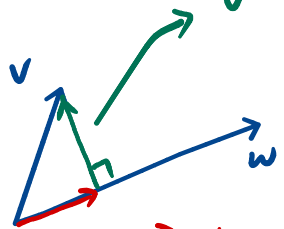
<u>Ans</u>: $\begin{pmatrix} \frac{1}{2} \\ -1/2 \\ 1/2 \end{pmatrix}$ #

4.4 Orthogonal Projections and Subspaces

§ Orthogonal projections

Recall :

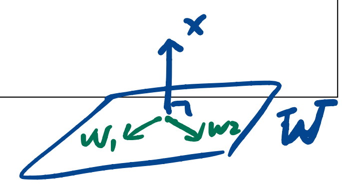
$v - \frac{\langle v, w\rangle}{\|w\|^2} w$, orthogonal to w.



$\frac{\langle v, w\rangle}{\|w\|^2} w$, orthogonal projection of x onto vector w.

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,

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



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 $y \in W$ can be written as

$$y = c_1 w_1 + \cdots + c_n w_n.$$

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

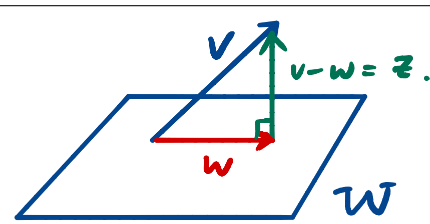
$$= c_1 \langle x, w_1 \rangle + \cdots + c_n \langle x, w_n \rangle$$

$$= 0.$$

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" $\mathbf{v}$.

Fact: 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

$$\mathbf{w} = \underline{\underline{c_1}}\mathbf{u}_1 + \cdots \underline{\underline{c_n}}\mathbf{u}_n,$$

where

$$\boxed{c_j = \langle \mathbf{v}, \mathbf{u}_j \rangle,} \quad j = 1, \cdots, n.$$

[To see this:]

(1) $w \in W$ can be written as linear combination of $\{u_1, \dots, u_n\}$, so

$$w = \underline{c_1} u_1 + \cdots + \underline{c_n} u_n.$$

$$\underline{\langle w, u_j \rangle} = \langle c_1 u_1 + \cdots + c_n u_n, u_j \rangle = c_j \langle u_j, u_j \rangle$$
$$= c_j \|u_j\|^2 = \underline{c_j}.$$

(2) Since w is orthogonal projection of v onto W,

$(v - w) \perp W$. $\quad u_j \in W$

so, $\langle v-w, u_j \rangle = 0 \Rightarrow \underline{\langle v, u_j \rangle} = \underbrace{\langle w, u_j \rangle}_{= c_j}$

- If $\{u_i\}$ is orthogonal, then $c_j = \dfrac{\langle v, u_j \rangle}{\|u_j\|^2}$. #

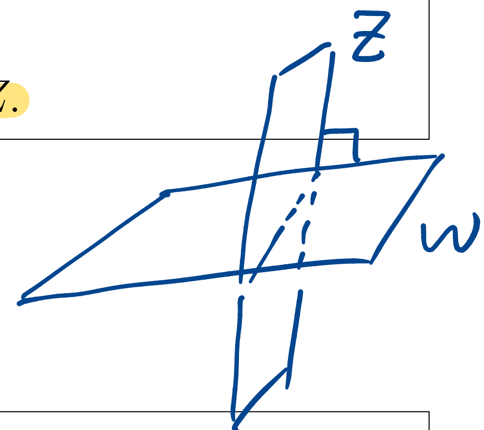
Remark: Thus, $\mathbf{v} - \underbrace{\sum_{k=1}^{n} \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k}_{= w}$ is orthogonal to W, that is,

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

§ 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,

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



Immediately, we also have

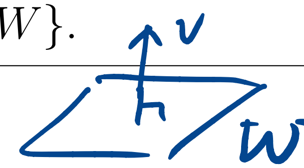
Fact: If $\{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$ span W and $\{\mathbf{z}_1, \cdots, \mathbf{z}_k\}$ span Z, then

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

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,

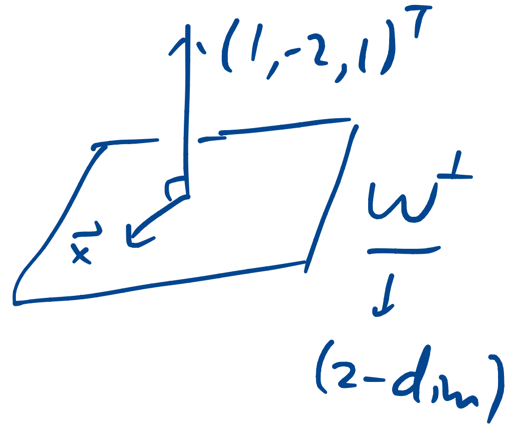
$$W^\perp = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ for all } \mathbf{w} \in W\}.$$



- 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 $\mathbf{0}$.

Example. Find $W^{\perp}$, the orthogonal complement to $W = \text{span}\{\mathbf{w}\}$ in $\mathbb{R}^3$, where

$$\mathbf{w} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$



$\vec{x} = (x, y, z)$ in $W^{\perp}$

$w^T \vec{x} = 0.$

$$\begin{pmatrix} \boxed{1} & -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0.$$

pivot

free variables.

$x = 2y - z$. $\quad W^{\perp} = \left\{ \begin{pmatrix} 2y - z \\ y \\ z \end{pmatrix} \middle| \; y, z \in \mathbb{R} \right\}$. #

Example. Suppose $W = \text{span}\{\mathbf{w}_1, \mathbf{w}_2\}$, where

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

To find $W^{\perp}$.

$w_1^T \vec{x} = 0$

$w_2^T \vec{x} = 0.$

$$\begin{bmatrix} w_1^T \\ w_2^T \end{bmatrix} \vec{x} = 0.$$

「 Continue on Monday 」