Lecture 16

Note: what you need to know about $\mathbb{C}$:

- How to divide complex numbers
- How to factor polynomials: if degree $\leq 2 \Rightarrow$ Use quadratic formula. If degree ≥ 3: look for integer solution $\lambda = n$ and divide the polynomial by $\lambda - n$.
- HW has to be tedious sometimes...

In-class exercises

Let $S = \text{Span}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right)$. Let $v = \begin{pmatrix}1\\2\\3\end{pmatrix}$. Find $v^{\parallel}$ and $v^{\perp}$.

Orthonormal basis: $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}$. Then $v^{\parallel} = \left(\begin{pmatrix}1\\0\\0\end{pmatrix}\cdot v\right)\cdot\begin{pmatrix}1\\0\\0\end{pmatrix} + \left(\begin{pmatrix}0\\1\\0\end{pmatrix}\cdot v\right)\cdot\begin{pmatrix}0\\1\\0\end{pmatrix} = 1\cdot\begin{pmatrix}1\\0\\0\end{pmatrix} + 2\cdot\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\2\\0\end{pmatrix}$.

$v^{\perp} = v - v^{\parallel} = \begin{pmatrix}1\\2\\3\end{pmatrix} - \begin{pmatrix}1\\2\\0\end{pmatrix} = \begin{pmatrix}0\\0\\3\end{pmatrix}$.

Recap:

- Length: $\|v\| = \sqrt{v_1^2 + \ldots + v_n^2}$
- Orthogonal vectors: $v \cdot w = 0$
- Unit vector: $\|u\| = 1$.
- Orthonormal vectors: $u_i \cdot u_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$
- Orthogonal projection: have a subspace $S \subseteq \mathbb{R}^n$ and an orthonormal basis $u_1, \ldots, u_s$.

  Then if $v \in \mathbb{R}^n$, the projection of v onto S is $v^{\parallel} = \text{proj}_S(v) = (u_1 \cdot v)u_1 + \ldots + (u_s \cdot v)u_s$.

  the perpendicular component is $v^{\perp} = v - v^{\parallel}$

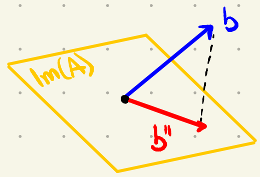
Today: the Gram-Schmidt process and QR factorization.

But first: <u>motivation for orthogonal projections</u>

Suppose you have an inconsistent system $Ax = b$.

Instead of calling it a day, we can find the vector $x_0 \in \mathbb{R}^n$ so that $\|Ax_0 - b\|$ is minimum.

Picture:



The vector $b'' = Ax_0$ is the closest vector to b inside the subspace $\mathrm{Im}(A)$.

This is sometimes called the "least squares solution" to the system $Ax=b$.

Last time we proved that projections are linear. Recall the example

Example 3 from lec 15 : Consider $S = \mathrm{Span}\left(\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right)$, and $v = \begin{pmatrix}1\\2\\3\end{pmatrix}$

Then
$$\left.\begin{aligned}
\mathrm{proj}_S\begin{pmatrix}1\\0\\0\end{pmatrix} &= (e_1\cdot u_1)\cdot u_1 + (e_1\cdot u_2)\cdot u_2 \\
&= \frac{1}{\sqrt{2}}\cdot u_1 + 0\cdot u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix} \\
\mathrm{proj}_S\begin{pmatrix}0\\1\\0\end{pmatrix} &= (e_2\cdot u_1)\cdot u_1 + (e_2\cdot u_2)\cdot u_2 \\
&= \frac{1}{\sqrt{2}} u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix} \\
\mathrm{proj}_S\begin{pmatrix}0\\0\\1\end{pmatrix} &= (e_1\cdot u_1)u_1 + (e_2\cdot u_2)u_2 = 0u_1 + u_2 = \begin{pmatrix}0\\0\\1\end{pmatrix}
\end{aligned}\right\}$$

proj_S is given by the matrix
$$\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Question: what is the matrix of proj_S?

Theorem 1: Let $S \subseteq \mathbb{R}^n$ be a subspace, and let $u_1, \dots, u_s \in S$ be an orthonormal basis for S. Let

$Q = \begin{pmatrix}| & & |\\ u_1 & \cdots & u_s\\ | & & |\end{pmatrix}$ ($n \times s$ matrix). Then the matrix of $\mathrm{proj}_S : \mathbb{R}^n \to \mathbb{R}^n$ is QQ^T.

Proof: We need to show that $\mathrm{proj}_S(e_i)$ is the ith column of QQ^T. Consider $u_1, \dots, u_s$ and extend them to a basis $\mathcal{B}$:

$u_1, \dots, u_s, v_{s+1}, \dots, v_n$ of $\mathbb{R}^n$. Then:

$$\begin{aligned}
\mathrm{proj}_S(e_i) &= (u_1\cdot e_i)u_1 + \dots + (u_s\cdot e_i)u_s \\
&= \begin{pmatrix}u_1\cdot e_i\\ \vdots\\ u_s\cdot e_i\\ 0\\ \vdots\\ 0\end{pmatrix}_{\mathcal{B}} \\
&= S_{\mathcal{B}\to\mathcal{E}}\begin{pmatrix}u_1\cdot e_i\\ \vdots\\ u_s\cdot e_i\\ 0\\ \vdots\\ 0\end{pmatrix} \\
&= \begin{pmatrix}| & & | & | & & |\\ u_1 & \cdots & u_s & v_{s+1} & \cdots & v_n\\ | & & | & | & & |\end{pmatrix}\begin{pmatrix}u_1\cdot e_i\\ \vdots\\ u_s\cdot e_i\\ 0\\ \vdots\\ 0\end{pmatrix} \\
&= \begin{pmatrix}\underbrace{\begin{matrix}| & & |\\ u_1 & \cdots & u_s\\ | & & |\end{matrix}}_{Q} & \begin{matrix}| & & |\\ v_{s+1} & \cdots & v_n\\ | & & |\end{matrix}\end{pmatrix}\begin{pmatrix}\left.\begin{matrix}- & u_1 & -\\ & \vdots & \\ - & u_s & -\end{matrix}\right\} Q^T\\ - & 0 & -\\ & \vdots & \\ - & 0 & -\end{pmatrix} e_i \\
&= QQ^T(e_i)
\end{aligned}$$

The question now is: how do we find orthonormal bases?

Answer: the Gram–Schmidt process.

Gram–Schmidt process:

Let $v_1, \dots, v_s$ be a basis of $S \subseteq \mathbb{R}^n$. Let $i = 1$ and $v_1^\perp = v_1$.

1. Normalize $v_i^\perp$: $u_i = \frac{v_i^\perp}{\|v_i^\perp\|}$. Let $i \to i+1$.

2. Make v_i orthogonal to $u_1, \dots, u_{i-1}$: $v_i^\perp = v_i - (u_1 \cdot v_i) u_1 - \dots - (u_{i-1} \cdot v_i) u_{i-1}$. Go back to 1.

Example 1: Let $S = \mathbb{R}^3$ with basis B: $v_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

1) $u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{1^2+1^2+1^2}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$

2) $v_2^\perp = v_2 - (u_1 \cdot v_2) u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{2}{\sqrt{3}} \cdot \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$

1) $u_2 = \frac{v_2^\perp}{\|v_2^\perp\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$

2) $v_3^\perp = v_3 - (u_1 \cdot v_3) u_1 - (u_2 \cdot v_3) u_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \frac{2}{\sqrt{3}} \cdot \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - \frac{3}{\sqrt{6}} \cdot \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 - \frac{2}{3} - \frac{1}{2} \\ 1 + \frac{2}{3} - \frac{1}{2} \\ 2 - \frac{2}{3} - 1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} -1 \\ 7 \\ 2 \end{pmatrix}$

1) $u_3 = \frac{1}{3\sqrt{6}} \begin{pmatrix} -1 \\ 7 \\ 2 \end{pmatrix}$.

Finally u_1, u_2, u_3 are an orthonormal basis of $S = \mathbb{R}^3$.

<u>QR factorization</u>

The Gram–Schmidt process produces a factorization of the matrix $A = \underbrace{\begin{pmatrix} | & & | \\ v_1 & \cdots & v_s \\ | & & | \end{pmatrix}}_{\text{l.i.}}$, as follows:

First, write $A = \underbrace{\begin{pmatrix} | & & | \\ v_1 & \cdots & v_s \\ | & & | \end{pmatrix}}_{\text{basis of } S}$.

Then Gram–Schmidt yields $v_1 = \|v_1\| u_1$,

$v_2 = \|v_2^\perp\| \cdot v_2^\perp + (u_1 \cdot v_2) u_1$

$v_3 = \|v_3^\perp\| \cdot v_3^\perp + (u_1 \cdot v_3) u_1 + (u_2 \cdot v_3) u_2$

$v_4 =$ etc.

In matrix form this says $A = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_s \\ | & & | \end{pmatrix} = \underbrace{\begin{pmatrix} | & & | \\ u_1 & \cdots & u_s \\ | & & | \end{pmatrix}}_{Q} \cdot \underbrace{\begin{pmatrix} \|v_1\| & u_1\cdot v_2 & u_1\cdot v_3 & \cdots & u_1\cdot v_s \\ 0 & \|v_2^\perp\| & u_2\cdot v_3 & & u_2\cdot v_s \\ 0 & 0 & \|v_3\| & \cdots & u_3\cdot v_s \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & \|v_s\| \end{pmatrix}}_{R \text{ (upper triangular)}}$

Definition 1: This is the <u>QR factorization</u> of A.

Example 1 (ctd'): $\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & -1/3\sqrt{3} \\ -1/\sqrt{3} & 2/\sqrt{6} & 7/3\sqrt{3} \\ 1/\sqrt{3} & 1/\sqrt{6} & 2/3\sqrt{3} \end{pmatrix} \underbrace{\begin{pmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ 0 & \frac{\sqrt{6}}{3} & \frac{3}{\sqrt{6}} \\ 0 & 0 & 3\sqrt{3} \end{pmatrix}}_{\text{we computed these when doing}}$ Gram-Schmidt.

<u>Orthogonal transformations</u>

Definition 2: An $n\times n$ matrix Q is orthogonal iff its columns consist of an orthonormal basis $\mathcal{B}: u_1, \dots, u_n$

$$Q = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix} \quad (= S_{\mathcal{B}\to e})$$

Recall that the transpose of a matrix Q is the matrix Q^T whose columns are the rows in Q

Fact: $(AB)^T = B^T A^T$ Proof: in-class exercise.

Example 2: $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$, $\begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 4 & -5 \\ 11 & -11 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 11 \\ -5 & -11 \end{pmatrix}$$

Definition 3: A linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ whose matrix is orthogonal is called an orthogonal transformation.

It turns out, finding the inverse of an orthogonal matrix is extremely easy:

Theorem 2: An $n\times n$ matrix is orthogonal iff $Q^TQ = I_n$.

Proof: Note that $Q^TQ = \begin{pmatrix} - u_1 - \\ \vdots \\ - u_n - \end{pmatrix}\begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix} = \begin{pmatrix} u_1\cdot u_1 & u_1\cdot u_2 & \cdots & u_1\cdot u_n \\ u_2\cdot u_1 & & & \vdots \\ \vdots & & & \\ u_n\cdot u_1 & & \cdots & u_n\cdot u_n \end{pmatrix}$

Evidently, this matrix is the identity if and only if $\mathcal{B}$ is orthonormal. $\square$

The associated transformation to Q satisfies an important property:

Theorem 3: Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation with associated matrix Q, and assume Q is orthogonal. Then, for any $v, w \in \mathbb{R}^n$, $T(v) \cdot T(w) = v \cdot w$

In particular, $\|T(v)\| = \|v\|$.

Proof: Take $v, w \in \mathbb{R}^n$. Then $T(v) \cdot T(w) = (Qv) \cdot (Qw) = (Qv)^T (Qw) = v^T Q^T Q w = v^T w = v \cdot w$.

- $\|T(v)\| = \|v\|$: In-class exercise. $\blacksquare$

In-class exercises:

1. Find the QR factorization of $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 2 \end{pmatrix}$

2. Let $S = \mathrm{Im}(A)$. Find the matrix of proj_S.

3. Prove:
   - If T is an orthogonal transformation, then $\|T(v)\| = \|v\|$ for all $v \in \mathbb{R}^n$ (Hint: use Thm 3).
   - If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then $(AB)^T = B^T A^T$.