

Wednesday, Sep 17, Lecture 9 : Inner products.

Last time...

- An $n \times n$ matrix M is diagonalizable iff there is an invertible matrix S st.

$$M = SDS^{-1} \quad (\text{or } = R^{-1}DR)$$

where D is diagonal.

- From perspective of linear operators: $T: V \rightarrow V$ is diagonalizable if there is a basis $\mathcal{B}_{\text{eig}} = \{v_1, \dots, v_n\}$ of eigenvectors of T :

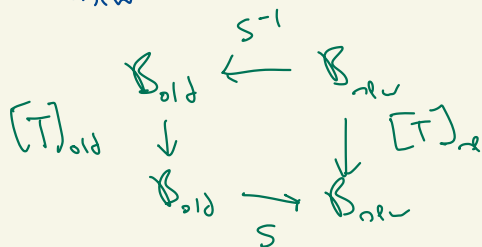
$$Tv_i = \lambda_i v_i$$

- The change-of-basis matrix S from $\mathcal{B}_{\text{old}} = \{e_1, \dots, e_n\}$ to $\mathcal{B}_{\text{new}} = \{f_1, \dots, f_n\}$ is

$$S = \begin{bmatrix} | & & | \\ [e_1]_{\text{new}} & \dots & [e_n]_{\text{new}} \\ | & & | \end{bmatrix}$$

- Take a linear map $T: V \rightarrow V$. If $[T]_{\text{old}}$ is the matrix of T in the old basis $\mathcal{B}_{\text{old}}$,

$$[T]_{\text{new}} = S [T]_{\text{old}} S^{-1}$$



Let's revisit our example from last time.

Ex: $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is the linear map $T(x, y) = (x+y, x+y)$

$$\mathcal{B}_{st} = \{e_1 = (1, 0), e_2 = (0, 1)\}$$

$$\mathcal{B}_{eig} = \{f_1 = (1, 1), f_2 = (1, -1)\}$$

Then S which takes $\mathcal{B}_{st} \rightarrow \mathcal{B}_{eig}$ is

$$S = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and S^{-1} (which takes $\mathcal{B}_{eig} \rightarrow \mathcal{B}_{st}$) is

$$S^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$\text{[check: } SS^{-1} = S^{-1}S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{.]}$$

Then

$$\begin{aligned} [T]_{eig} &= S [T]_{st} S^{-1} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Observation: (Diagonalizable matrices and linear operators)

Linear map
 $T: V \rightarrow V$ is
diagonalizable

$\Leftrightarrow$ $n \times n$ matrix $[T]_{\text{orig}}$ in any basis $\mathcal{B}_{\text{orig}}$
diagonalizable:

$$[T]_{\text{orig}} = S [T]_{\text{eig}} S^{-1}$$

$\uparrow$
diagonal!

Remember, if $\mathcal{B}_{\text{eig}} = \{v_1, \dots, v_n\}$ is an eigenbasis, i.e.
 $T v_i = \lambda_i v_i \quad i=1, \dots, n,$

This is the same as

$$[T]_{\text{eig}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

The change-of-basis goes

so

$$\mathcal{B}_{\text{eig}} \xrightarrow{S} \mathcal{B}_{\text{orig}} \xrightarrow{S^{-1}} \mathcal{B}_{\text{eig}}$$

And the columns of S are eigenvectors v_i of T (in orig. basis)

$$S = \begin{bmatrix} | & & | \\ [v_1]_{\text{orig}} & \dots & [v_n]_{\text{orig}} \\ | & & | \end{bmatrix}$$

Related: see problem 4 HW4.

Warning: not every matrix is diagonalizable,

e.g.

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Here, $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector,

but there are no other linearly
indip eigenvectors:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ y \end{bmatrix}$$

if $y \neq 0$, there
is no λ such that
 $x = \lambda(x+y)$
 $y = \lambda y$

- A big question: when are linear maps diagonalizable?
- Another: in physics and in life we have notions of length and angle. where is it now?
- The answers to both lie in Hilbert Spaces, vector spaces with an inner product.

Inner Products

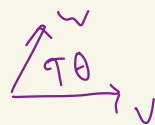
We want notions of length and angle.

Ex: $V = \mathbb{R}^n$. We have dot product: if $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$,

$$\langle v, w \rangle = v \cdot w = v_1 w_1 + v_2 w_2 + \dots + v_n w_n$$

$$= \|v\| \|w\| \cos \theta$$

$\uparrow$ $\uparrow$
 length angle
 of v
 $= \sqrt{\langle v, v \rangle}$



HW: $V = C([0,1])$, continuous functions,

$$\langle v, w \rangle = \int_0^1 v(x) w(x) dx$$

Hilbert Space = vector space + inner product

Def: Inner product

Let V be a ^(complex) real vector space. Let $\alpha \in \mathbb{R}$, $v, w \in V$.

Then an inner product $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ ($V \times V \rightarrow \mathbb{C}$) must satisfy:

① (Positive definite) $\langle v, v \rangle \geq 0$, and if $\langle v, v \rangle = 0$ then $v = 0$.

② (Linearity in second arg) $\langle v, \alpha w + u \rangle = \alpha \langle v, w \rangle + \langle v, u \rangle$.

③ (Symmetric) $\langle v, w \rangle = \langle w, v \rangle$ ($\langle v, w \rangle = \overline{\langle w, v \rangle}$)

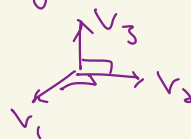
Def: orthonormal

A set of vectors $\{v_1, \dots, v_n\}$ is called orthogonal if

$$\langle v_i, v_j \rangle = 0 \quad i \neq j \quad \leftarrow \text{"all form right angles"}$$

and orthonormal if they also satisfy

$$\|v_i\|^2 = \langle v_i, v_i \rangle = 1. \quad \leftarrow \text{all have same length}$$



Ex:

• standard basis $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ of $\mathbb{R}^2$ (or $\mathbb{C}^2$)

• The basis $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is orthogonal:

$$\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle = 1 \cdot 1 + 1 \cdot (-1) = 0$$

But not orthonormal, since $\|\begin{pmatrix} 1 \\ 1 \end{pmatrix}\| = \sqrt{2}$.

This is easily fixed:

$$\tilde{v} = \frac{v}{\|v\|} \quad \text{always has norm 1,}$$

$$\text{since } \|\tilde{v}\| = \left\| \frac{v}{\|v\|} \right\| = \frac{\|v\|}{\|v\|} = 1.$$

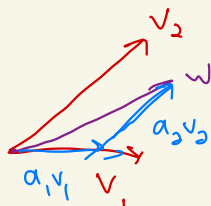
$\Rightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an orthonormal basis

(More generally: the Gram-Schmidt algorithm).

- One reason orthonormal bases ("onb") are so nice:
calculating coordinates is much easier

Ex:

. Not an onb: $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

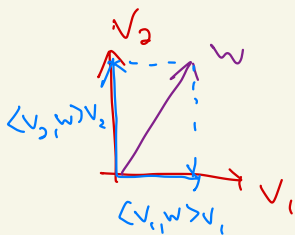


To find coordinates for w , need to solve

$$w = a_1 v_1 + a_2 v_2$$

This is geometrically annoying, and algebraically requires you to solve a system of equations.

. Onb: $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$



$$w = \langle v_1, w \rangle v_1 + \langle v_2, w \rangle v_2$$

Geometrically clear and algebraically easy.

Why it works:

if $B_{\text{onb}} = \{v_1, \dots, v_n\}$ is orthonormal and $w \in V$,

then

$$w = w_1 v_1 + w_2 v_2 + \dots + w_n v_n.$$

Pick a basis vector v_i . Then

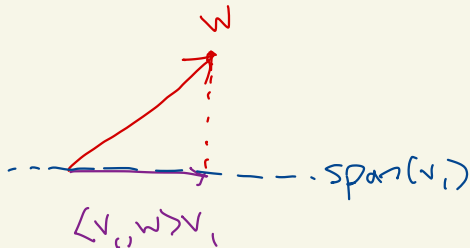
$$\begin{aligned}\langle v_i, w \rangle &= \langle v_i, w_1 v_1 + w_2 v_2 + \dots + w_n v_n \rangle \\ &= w_1 \langle v_i, v_1 \rangle + w_2 \langle v_i, v_2 \rangle + \dots + w_i \langle v_i, v_i \rangle + \dots + w_n \langle v_i, v_n \rangle \\ &= w_i\end{aligned}$$

Thus we have coordinates:

$$w = \underbrace{\langle v_1, w \rangle}_{\text{or}} v_1 + \dots + \langle v_i, w \rangle v_i + \dots + \langle v_n, w \rangle v_n,$$

$$[w]_{\text{onb}} = \begin{bmatrix} \langle v_1, w \rangle \\ \langle v_2, w \rangle \\ \vdots \\ \langle v_n, w \rangle \end{bmatrix}$$

Remember, if $\|v_i\| = 1$,
this is the overlap:



Any basis may be converted to an orthonormal one

$$\{v_1, \dots, v_n\} \xrightarrow{\text{G.S.}} \{\tilde{v}_1, \dots, \tilde{v}_n\}$$

by the Gram-Schmidt algorithm:

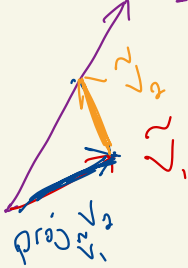
skip lecture $\left[\begin{array}{l} \text{Ex: Convert } v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ into an orb.} \end{array} \right.$

$$\begin{array}{c} \nearrow \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \nearrow \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{array}$$

• Normalize:

$$\tilde{v}_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\tilde{v}_1 \rightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad v_2$$



• Remove overlap:

$$\begin{aligned} v_2^{(0)} &= v_2 - \frac{\langle v_2, \tilde{v}_1 \rangle}{\|\tilde{v}_1\|^2} \tilde{v}_1 \\ &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot \frac{3}{1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 - 3/2 \\ 1 - 3/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned}$$

Normalize:

$$\tilde{v}_2 = \frac{v_2^{(0)}}{\|v_2^{(0)}\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

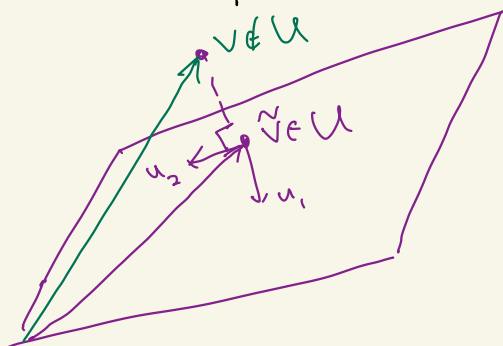
check:

$$\langle \tilde{v}_1, \tilde{v}_2 \rangle = \frac{1}{2} (1 \cdot 1 + (-1) \cdot 1) = 0$$

$\left[\text{and } \|\tilde{v}_1\| = \|\tilde{v}_2\| = 1 \right.$

Orthogonal Projections

Where orthogonality really shines: Suppose we have a subspace $U \subseteq V$ and a point $v \notin U$.



How can we find the closest point $\tilde{v} \in U$?
Really need orthogonality here!

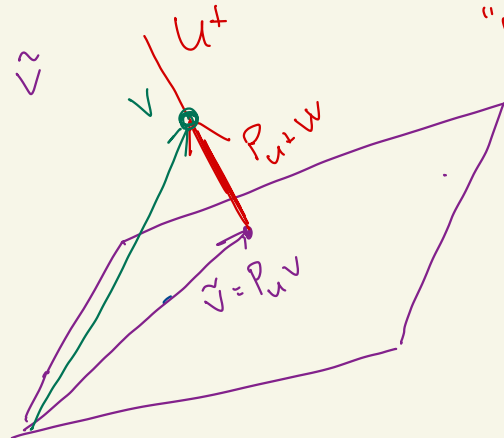
Suppose $\{u_1, \dots, u_m, w_{m+1}, \dots, w_n\}$ is an orb of V and $\{u_1, \dots, u_m\}$ is an orb of U .

Then

$$v = \underbrace{\langle u_1, v \rangle u_1 + \dots + \langle u_m, v \rangle u_m}_{\text{projection onto } U} + \underbrace{\langle w_{m+1}, v \rangle w_{m+1} + \dots + \langle w_n, v \rangle w_n}_{\text{projection onto } U^\perp}$$

$$= \underbrace{P_U v}_{\tilde{v}} + P_{U^\perp} v$$

orthogonal complement
"U-perp"



Many applications! Projections give best* approximations in lower dimensional spaces

Ex: HW 4, Problem 5

"Find the degree 2 polynomial which best approximates* $\cos(2\pi x)$ on $[0,1]$ "

* this notion depends on $\langle, \rangle$.

Def: Orthogonal Projection

Let V be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$, and let U be a subspace of V .

The orthogonal projection $P: V \rightarrow V$ is the linear map satisfying

① $\text{ran}(P) = U$ (P projects onto U)

② $P^2 = P$ (P acts as identity on U : if $u \in U$, $Pu = u$).

③ $\langle Pv, w \rangle = \langle v, Pw \rangle$ for all $v, w \in V$. (P is self-adjoint)

↑ Two special cases of ③, to be discussed:

- $P = P^T$ (real vector spaces)

- $P = P^+$ (complex vector spaces)