5.1   Inner Product Spaces

- Recall that for $z = x + iy \in \mathbb{C}$, $\bar{z} = x - iy$ and $|z|^2 = \bar{z}z = x^2 + y^2$.

- For $\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$ we define the <u>inner product</u> of $\vec{z}$ and $\vec{w}$ by

$$\langle \vec{z}, \vec{w} \rangle := \sum_{i=1}^{n} \bar{w}_i z_i$$

- For $\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n$ we define the <u>norm</u> of $\vec{z}$ by

$$\|\vec{z}\| = \langle \vec{z}, \vec{z} \rangle^{1/2} = \left(\sum_{i=1}^{n} \bar{z}_i z_i \right)^{1/2} = \left(\sum_{i=1}^{n} |z_i|^2 \right)^{1/2}$$

$\boxed{\text{EX}}$ Let $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$. Then

$$\|\vec{x}\| = \sqrt{|x_1|^2 + |x_2|^2} = \sqrt{x_1^2 + x_2^2}$$

<u>length of vector</u>
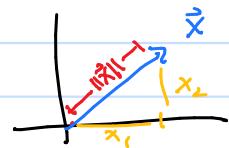
$\boxed{\text{Def}}$ For $A \in M_{m \times n}(\mathbb{C})$, the <u>adjoint</u> of A is the matrix $A^* \in M_{n \times m}(\mathbb{C})$ with entries $(A^*)_{i,j} = \overline{(A)_{j,i}}$.

- For $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$, if we think of $\vec{w} \in M_{n \times 1}(\mathbb{C})$, then

$$\vec{w}^* = (\bar{w}_1, \ldots, \bar{w}_n) \in M_{1 \times n}(\mathbb{C})$$

and for $\vec{z} \in \mathbb{C}^n$

← matrix multiplication.

$$\langle \vec{z}, \vec{w} \rangle = \bar{w}_1 z_1 + \cdots + \bar{w}_n z_n = \vec{w}^* \vec{z}$$

$\boxed{\text{Rem}}$ Note that the inner product satisfies the following properties

① $\langle \vec{w}, \vec{z} \rangle = \overline{\langle \vec{z}, \vec{w} \rangle}$

② $\langle \alpha \vec{z}_1 + \beta \vec{z}_2, \vec{w} \rangle = \alpha \langle \vec{z}_1, \vec{w} \rangle + \beta \langle \vec{z}_2, \vec{w} \rangle$ for all $\alpha, \beta \in \mathbb{C}$

③ $\langle \vec{z}, \vec{z} \rangle = \|\vec{z}\|^2 \geq 0$

④ $\langle \vec{z}, \vec{z} \rangle = 0 \implies \vec{z} = \vec{0}$.

$\boxed{\text{Def}}$ Let V be vector space. A <u>inner product</u> on V is a map that assigns to each pair of vectors $\vec{v}, \vec{w} \in V$ a scalar <u>denoted</u> $\langle \vec{v}, \vec{w} \rangle$ with the following properties:

① Conjugate symmetry: $\langle \vec{w}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{w} \rangle}$ for all $\vec{v}, \vec{w} \in V$.

    (If $\mathbb{F} = \mathbb{R}$, then this replaced with symmetry: $\langle \vec{w}, \vec{v} \rangle = \langle \vec{v}, \vec{w} \rangle$ for all $\vec{v}, \vec{w} \in V$.)

② Linearity: $\langle \alpha \vec{u} + \beta \vec{v}, \vec{w} \rangle = \alpha \langle \vec{u}, \vec{w} \rangle + \beta \langle \vec{v}, \vec{w} \rangle$ for scalars α, β and all $\vec{u}, \vec{v}, \vec{w} \in V$.

③ Non-negativity: $\langle \vec{v}, \vec{v} \rangle \geq 0$ for all $\vec{v} \in V$.

④ Non-degeneracy: $\langle \vec{v}, \vec{v} \rangle = 0$ if and only if $\vec{v} = \vec{0}$.

We call V together with the map $\langle \cdot, \cdot \rangle$ an <u>inner product space</u>. For each $\vec{v} \in V$, the <u>norm</u> of $\vec{v}$ is the quantity:

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}.$$

Ex ① $\mathbb{C}^n$ (and $\mathbb{R}^n$) is an inner product space with inner product
$$\langle \vec{z}, \vec{\omega} \rangle = \sum_{i=1}^{n} \overline{\omega_i}\, z_i.$$

② Define
$$\ell^2(\mathbb{N}) = \left\{ (a_n)_{n \in \mathbb{N}} : a_n \in \mathbb{C}, \ \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}$$

Then this is a vector space with operations
$$(a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}} = (a_n + b_n)_{n \in \mathbb{N}}$$
$$\alpha (a_n)_{n \in \mathbb{N}} = (\alpha\, a_n)_{n \in \mathbb{N}}$$

Moreover, it has an inner product:
$$\langle (a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \rangle_2 := \sum_{n=1}^{\infty} \overline{b_n}\, a_n,$$
but one has to justify why the series on the right converges. Note that
$$\| (a_n)_{n \in \mathbb{N}} \|_2 := \left(\sum_{n=1}^{\infty} |a_n|^2 \right)^{1/2}.$$

③ Let $C[0,1]$ be the vector space of cts functions $f: [0,1] \to \mathbb{F}$ with operations
$$(f + g)(t) = f(t) + g(t)$$
$$(\alpha f)(t) = \alpha\, f(t)$$

This is an inner product space with inner product:
$$\langle f, g \rangle := \int_0^1 f(t)\, \overline{g(t)}\, dt.$$
Note that for each $n \in \mathbb{N}$, $P_n \subseteq C[0,1]$.

④ $M_{m \times n}$ can be made into an inner product space:
$$\langle A, B \rangle_2 := \operatorname{Tr}(B^* A) = \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{B_{ij}}\, A_{ij}$$

Note that
$$\| A \|_2 = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |A_{ij}|^2 \right)^{1/2}.$$

Properties of Inner Product Spaces

Fix an inner product space $(V, \langle \cdot, \cdot \rangle)$. Note that ① and ② imply:
$$\langle \vec{u}, \alpha \vec{v} + \beta \vec{\omega} \rangle = \overline{\langle \alpha \vec{v} + \beta \vec{\omega}, \vec{u} \rangle} = \overline{\alpha \langle \vec{v}, \vec{u} \rangle + \beta \langle \vec{\omega}, \vec{u} \rangle}$$
$$= \overline{\alpha} \langle \vec{v}, \vec{u} \rangle + \overline{\beta} \langle \vec{\omega}, \vec{u} \rangle = \overline{\alpha} \langle \vec{u}, \vec{v} \rangle + \overline{\beta} \langle \vec{u}, \vec{\omega} \rangle.$$

Thus the inner product is conjugate linear in the second coordinate.
$$2' \qquad \langle \vec{u}, \alpha \vec{v} + \beta \vec{\omega} \rangle = \overline{\alpha} \langle \vec{u}, \vec{v} \rangle + \overline{\beta} \langle \vec{u}, \vec{\omega} \rangle$$

② also implies (Exercise)
$$* \qquad \langle \vec{v}, \vec{0} \rangle = \langle \vec{0}, \vec{v} \rangle = 0 \qquad \forall \vec{v} \in V.$$

Lemma For $\vec{v} \in V$, $\vec{v} = \vec{0}$ if and only if
$$\langle \vec{v}, \vec{\omega} \rangle = 0 \qquad \forall \vec{\omega} \in V.$$

Proof ($\Longrightarrow$) This follows from $(*)$.

($\Longleftarrow$) Note $\| \vec{v} \| = (\langle \vec{v}, \vec{v} \rangle)^{1/2} = 0$. So $\vec{v} = \vec{0}$ by ④

Cor [5.2] For $\vec{x}, \vec{y} \in V$, $\vec{x} = \vec{y}$ if and only if
$$\langle \vec{x}, \vec{w} \rangle = \langle \vec{y}, \vec{w} \rangle \qquad \forall \vec{w} \in V.$$

Moreover, if $A, B : U \to V$ are linear transformations, then $A = B$ if and only if
$$\langle A\vec{u}, \vec{w} \rangle = \langle B\vec{u}, \vec{w} \rangle \qquad \forall \vec{u} \in U, \vec{w} \in W.$$

Proof For the first statement, apply the previous lemma to $\vec{v} := \vec{x} - \vec{y}$ and use linearity. For the second statement, the only if direction ($\Rightarrow$) is obvious. For the other direction, fix $\vec{u} \in U$. Then the first statement implies $A\vec{u} = B\vec{u}$. Since $\vec{u} \in U$ was arbitrary, we have $A = B$. $\square$

Thm [5.3] (Cauchy-Schwarz Inequality)
Let V be an inner product space. Then for any $\vec{x}, \vec{y} \in V$
$$|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \, \|\vec{y}\|.$$

Proof Let $t \in \mathbb{C}$ and consider
$$0 \leq \|\vec{x} - t\vec{y}\|^2 = \langle \vec{x} - t\vec{y}, \vec{x} - t\vec{y} \rangle$$
$$= \langle \vec{x}, \vec{x} - t\vec{y} \rangle - t \langle \vec{y}, \vec{x} - t\vec{y} \rangle$$
$$= \langle \vec{x}, \vec{x} \rangle - \bar{t} \langle \vec{x}, \vec{y} \rangle - t \langle \vec{y}, \vec{x} \rangle + |t|^2 \langle \vec{y}, \vec{y} \rangle$$
$$= \|\vec{x}\|^2 - \bar{t} \langle \vec{x}, \vec{y} \rangle - t \overline{\langle \vec{x}, \vec{y} \rangle} + |t|^2 \|\vec{y}\|^2$$

Set $t := \dfrac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|^2}$. Then the above becomes
$$0 \leq \cdots = \|\vec{x}\|^2 - \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^2} - \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^2} + \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^4} \cdot \|\vec{y}\|^2$$
$$= \|\vec{x}\|^2 - \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^2}$$

Thus
$$\frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\|\vec{y}\|^2} \leq \|\vec{x}\|^2 \implies |\langle \vec{x}, \vec{y} \rangle|^2 \leq \|\vec{x}\|^2 \cdot \|\vec{y}\|^2. \qquad \square$$

EX For $V = \mathbb{R}^2$, you might recall
$$\langle \vec{v}, \vec{w} \rangle = \cos\theta \, \|\vec{v}\| \cdot \|\vec{w}\|$$
where θ is the angle between $\vec{v}$ and $\vec{w}$. Since $-1 \leq \cos\theta \leq 1$, we have $|\cos\theta| \leq 1$, so the above formula implies the Cauchy-Schwarz inequality:
$$|\langle \vec{v}, \vec{w} \rangle| = |\cos\theta| \cdot \|\vec{v}\| \cdot \|\vec{w}\| \leq \|\vec{v}\| \, \|\vec{w}\|. \qquad \square$$

Thm [5.4] Let V be an inner product space, and let $\vec{x}, \vec{y} \in V$.
  ① (Triangle Inequality) $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|.$

  ② (Polarization Identity) just these if V is real vector space
  $$\langle \vec{x}, \vec{y} \rangle = \frac{1}{4}\left(\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2 + i\|\vec{x} + i\vec{y}\|^2 - i\|\vec{x} - i\vec{y}\|^2 \right) = \frac{1}{4} \sum_{k=1}^{4} i^k \|\vec{x} + i^k \vec{y}\|^2$$

  ③ (Parallelogram Identity) $\|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2 = 2(\|\vec{x}\|^2 + \|\vec{y}\|^2)$

Proof All three parts are proven by direct computation.

① We compute

$$\|\vec{x} + \vec{y}\|^2 = \langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle = \|\vec{x}\|^2 + \overbrace{\langle \vec{x}, \vec{y} \rangle + \overline{\langle \vec{x}, \vec{y} \rangle}}^{= 2 \cdot \text{real part of } \langle \vec{x}, \vec{y} \rangle} + \|\vec{y}\|^2$$

$$\leq \|x\|^2 + 2|\langle \vec{x}, \vec{y} \rangle| + \|\vec{y}\|^2$$

$$\text{Cauchy-Schwarz} \leq \|x\|^2 + 2\|\vec{x}\| \cdot \|\vec{y}\| + \|\vec{y}\|^2 = \left(\|\vec{x}\| + \|\vec{y}\|\right)^2$$

② We first compute

$$\|\vec{x} + \vec{y}\|^2 + i\|\vec{x} + i\vec{y}\|^2 = \|\vec{x}\|^2 + \langle \vec{x}, \vec{y} \rangle + \overline{\langle \vec{x}, \vec{y} \rangle} + \|\vec{y}\|^2 + i\left(\|\vec{x}\|^2 - i\langle \vec{x}, \vec{y} \rangle + i\overline{\langle \vec{x}, \vec{y} \rangle} + \|\vec{y}\|^2\right)$$

$$= (1+i)\left(\|\vec{x}\|^2 + \|\vec{y}\|^2\right) + 2\langle \vec{x}, \vec{y} \rangle$$

Swapping $-\vec{y}$ for $\vec{y}$ yields

$$\|\vec{x} - \vec{y}\|^2 + i\|\vec{x} - i\vec{y}\|^2 = (1+i)\left(\|\vec{x}\|^2 + \|\vec{y}\|^2\right) - 2\langle \vec{x}, \vec{y} \rangle.$$

Thus

$$\sum_{k=1}^{4} (i)^k \|\vec{x} + (i)^k \vec{y}\|^2 = 4\langle \vec{x}, \vec{y} \rangle.$$

③ We compute

$$\|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2 = \|\vec{x}\|^2 + \langle \vec{x}, \vec{y} \rangle + \overline{\langle \vec{x}, \vec{y} \rangle} + \|\vec{y}\|^2 + \|\vec{x}\|^2 - \langle \vec{x}, \vec{y} \rangle - \overline{\langle \vec{x}, \vec{y} \rangle} + \|\vec{y}\|^2$$

$$= 2\left(\|\vec{x}\|^2 + \|\vec{y}\|^2\right) \qquad \square$$

EX For $V = \mathbb{R}^2$, ① says the length of $\vec{x} + \vec{y}$ is at most the length of $\vec{x}$ plus the length of $\vec{y}$. In other words, its shorter to travel straight along $\vec{x} + \vec{y}$, rather than along $\vec{x}$ then along $\vec{y}$.
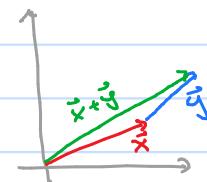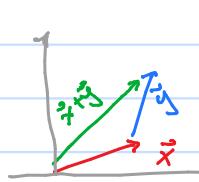
③ is a fact from geometry about parallelograms: the sum of the squares of the diagonals equals the sum of the squares of all four side lengths. $\square$
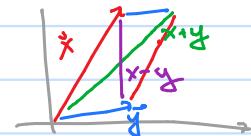
12/2

Normed Vector Spaces

- We have seen that all inner products give rise to norms, but these are objects that can be studied independently.

Def For a vector space V, a norm is a map from V to $\mathbb{R}$ satisfying:

① (Homogeneity) $\|\alpha \vec{v}\| = |\alpha| \|\vec{v}\|$ for all scalars α and $\vec{v} \in V$.

② (Triangle Inequality) $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$ for all $\vec{u}, \vec{v} \in V$.

③ (Non-negativity) $\|\vec{v}\| \geq 0$ for all $\vec{v} \in V$

④ (Non-degeneracy) $\|\vec{v}\| = 0$ if and only if $\vec{v} = \vec{0}$.

We call V along with the map $\|\cdot\|$ a <u>normed vector space</u>.

• Clearly every inner product space is a normed vector space, but the converse is not true.

EX. Fix $1 \le p < \infty$. The following norms come from inner products iff $p = 2$.

① For $\vec{x} = (x_1, ..., x_n)^T \in \mathbb{F}^n$ define
$$\| \vec{x} \|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \leftarrow \text{to ensure homogeneity}$$

② Define
$$\ell^p(\mathbb{N}) := \left\{ (a_n)_{n \in \mathbb{N}} : a_n \in \mathbb{C}, \quad \sum_{n=1}^{\infty} |a_n|^p < \infty \right\}$$
with
$$\| (a_n)_{n \in \mathbb{N}} \|_p := \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{1/p}$$

③ For $f \in C[0,1]$ define
$$\| f \|_p := \left(\int_0^\infty |f(t)|^p dt \right)^{1/p}$$

④ For $A \in M_{m \times n}$ define
$$\| A \|_p := \text{Tr} \left((A^* A)^{p/2} \right)^{1/p}$$

One can also make sense of $p = \infty$ for each of the above. For example
$$\| \vec{x} \|_\infty = \max_{1 \le i \le n} |x_i| \qquad \text{and} \qquad \| f \|_\infty = \max_{0 \le t \le 1} |f(t)|.$$

1. One way to tell if a norm comes from an inner product is the following theorem.

Thm 5.5 Let V be a normed vector space with norm $\| \cdot \|$. Then the norm comes from an inner product if and only if it satisfies the parallelogram identity:
$$\| \vec{x} + \vec{y} \|^2 + \| \vec{x} - \vec{y} \|^2 = 2 (\| \vec{x} \|^2 + \| \vec{y} \|^2) \qquad \forall \vec{x}, \vec{y} \in V.$$

Proof ($\Longrightarrow$) This is ③ from the previous theorem. (Thm 5.4)
$\quad (\Longleftarrow)$ Homework 10.

Prop 5.6 For $A \in M_{m \times n}$, $\vec{v} \in \mathbb{F}^n$, $\vec{w} \in \mathbb{F}^m$
$$\langle A \vec{v}, \vec{w} \rangle = \langle \vec{v}, A^* \vec{w} \rangle$$

Proof Treating $\vec{v}$ and $\vec{w}$ as matrices we have:
$$\langle \vec{v}, A^* \vec{w} \rangle = (A^* \vec{w})^* \vec{v} = \vec{w}^* A \vec{v} = \langle A \vec{v}, \vec{w} \rangle.$$

Def In an inner product space V, we say $\vec{v}, \vec{w} \in V$ are orthogonal if $\langle \vec{v}, \vec{w} \rangle = 0$, in which case we write $\vec{v} \perp \vec{w}$.

EX ① In $\mathbb{F}^3$,
$$\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ a \\ b \end{pmatrix} \right\rangle = 1 \cdot 0 + 0 \cdot a + 0 \cdot b = 0$$
In particular $\vec{e}_1 \perp \vec{e}_2$ and $\vec{e}_1 \perp \vec{e}_3$

② For the std. basis $\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{F}^n$, $\vec{e}_i \perp \vec{e}_j$ whenever $i \neq j$.

③ $C[-1,1]$ has inner product
$$\langle f, g \rangle_2 = \int_{-1}^{1} f(t) \overline{g(t)} \, dt$$
Let $f(t) = t^{2n}$, $g(t) = t^{2m+1}$. Then
$$\langle f, g \rangle_2 = \int_{-1}^{1} t^{2n} t^{2m+1} dt = \int_{-1}^{1} t^{2(n+m)+1} dt = 0$$

so $f \perp g$.

- Observe that $\vec{v} \perp \vec{v} \iff \langle \vec{v}, \vec{v} \rangle = 0 \iff \vec{v} = \vec{0}$.
- If $\vec{v} \perp \vec{w}$ then
$$\| \vec{v} + \vec{w} \|^2 = \| \vec{v} \|^2 + \| \vec{w} \|^2 \qquad (\text{Pythagorean Identity})$$

Indeed:
$$\| \vec{v} + \vec{w} \|^2 = \langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \rangle = \| \vec{v} \|^2 + \langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{v} \rangle + \| \vec{w} \|^2 = \| \vec{v} \|^2 + \| \vec{w} \|^2$$

Def Let V be an inner product space. We say $\vec{v} \in V$ is orthogonal to a subspace $E \subset V$ if $\vec{v} \perp \vec{w}$ for all $\vec{w} \in E$. We write $\vec{v} \perp E$.
We say two subspaces $E, F \subset V$ are orthogonal if $\vec{v} \perp \vec{w}$ for all $\vec{v} \in E$ and $\vec{w} \in F$.

EX Let $A \in M_{m \times n}(\mathbb{C})$. Then $\text{Ker}(A) \perp \text{Ran}(A^*)$. Indeed, let $\vec{v} \in \text{Ker}(A)$ and $\vec{w} \in \text{Ran}(A^*)$. Then $\vec{w} = A^* \vec{u}$ for some $\vec{u} \in \mathbb{C}^m$. Now
$$\langle \vec{v}, \vec{w} \rangle = \vec{w}^\# \vec{v} = (A^* \vec{u})^* \vec{v} = (\vec{u}^* A) \vec{v} = \vec{u}^* (A \vec{v}) = \vec{u}^* \vec{0} = 0.$$
Thus $\vec{v} \perp \vec{w}$.

Lemma 5.7 Let $\vec{v}_1, \ldots, \vec{v}_r \in V$ and let $E = \text{span}\{\vec{v}_1, \ldots, \vec{v}_r\}$. Then for $\vec{v} \in V$, we have $\vec{v} \perp E$ if and only if $\vec{v} \perp \vec{v}_k$ for each $k = 1, \ldots, r$.

Proof ($\implies$) Since $\vec{v}_1, \ldots, \vec{v}_r \in E$, this follows by definition of $\vec{v} \perp E$.
($\impliedby$) Suppose $\vec{v} \perp \vec{v}_k$ for each $k = 1, \ldots, r$. Let $\vec{w} \in E$, then
$$\vec{w} = \alpha_1 \vec{v}_1 + \cdots + \alpha_r \vec{v}_r$$

for some scalars $\alpha_1, \ldots, \alpha_r$. We have

$$\langle \vec{v}, \vec{w} \rangle = \langle \vec{v}, \alpha_1 \vec{v}_1 + \cdots + \alpha_r \vec{v}_r \rangle = \overline{\alpha_1} \langle \vec{v}, \vec{v}_1 \rangle + \cdots + \overline{\alpha_r} \langle \vec{v}, \vec{v}_r \rangle = 0$$

So $\vec{v} \perp \vec{w}$, and since $\vec{w} \in E$ was arbitrary we have $\vec{v} \perp E$. $\square$

Def In an inner product space V, a system $\vec{v}_1, .., \vec{v}_n \in V$ is called <u>orthogonal</u> if $\vec{v}_i \perp \vec{v}_j$ whenever $i \neq j$. The system is said to be <u>orthonormal</u> if it is orthogonal and $\|\vec{v}_i\| = 1$ for each $i = 1, .., n$.

EX In $\mathbb{R}^n$, $\hat{e}_1, .., \hat{e}_n$ is orthonormal. $\square$

Lemma 5.8 (Generalized Pythagorean Identity) Let $\vec{v}_1, ..., \vec{v}_n$ be an orthogonal system. Then

$$\| \alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n \|^2 = |\alpha_1|^2 \|\vec{v}_1\|^2 + \cdots + |\alpha_n|^2 \|\vec{v}_n\|^2$$

for any scalars $\alpha_1, .., \alpha_n$.

Proof We compute

$$\| \sum_{j=1}^{n} \alpha_j \vec{v}_j \|^2 = \langle \sum_{j=1}^{n} \alpha_j \vec{v}_j , \sum_{k=1}^{n} \alpha_k \vec{v}_k \rangle$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_j \overline{\alpha_k} \langle \vec{v}_j, \vec{v}_k \rangle$$

$$= \sum_{j=1}^{n} \alpha_j \overline{\alpha_j} \langle \vec{v}_j, \vec{v}_j \rangle = \sum_{j=1}^{n} |\alpha_j|^2 \|\vec{v}_j\|^2. \quad \square$$

Cor 5.9 Any orthogonal system $\vec{v}_1, .., \vec{v}_n$ of <u>non-zero</u> vectors is linearly independent.

Proof Suppose

$$\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n = \vec{0}$$

for some scalars $\alpha_1, ..., \alpha_n$. Then using the lemma we have

$$0 = \|\vec{0}\|^2 = \| \alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n \|^2 = |\alpha_1|^2 \|\vec{v}_1\|^2 + \cdots + |\alpha_n|^2 \|\vec{v}_n\|^2$$

Note $\|\vec{v}_j\| \neq 0$ since $\vec{v}_j \neq \vec{0}$ for each $j = 1, ..., n$. So the only way the sum of non-negative numbers on the right equals zero is if $|\alpha_1|^2 = \cdots = |\alpha_n|^2 = 0 \implies \alpha_1 = \cdots = \alpha_n = 0$. Thus $\vec{v}_1, ..., \vec{v}_n$ are linearly independent. $\square$

Ex In $C[-1, 1]$, consider

$$f_n(t) := e^{i \pi n t} \qquad n \in \mathbb{Z}$$

We claim $f_n \perp f_m$ for $n \neq m$. **Fact** $e^{ix} = \cos(x) + i \sin(x)$ for $x \in \mathbb{R}$. Thus

$$\langle f_n, f_m \rangle_2 = \int_{-1}^{1} f_n(t) \overline{f_m(t)} \, dt = \int_{-1}^{1} e^{i \pi n t} e^{-i \pi m t} \, dt = \int_{-1}^{1} e^{i \pi (n-m) t} \, dt$$

$$= \int_{-1}^{1} \cos(\pi(n-m)t) + i \sin(\pi(n-m)t) \, dt$$

$$\overset{n-m \neq 0}{=} \left[\frac{\sin(\pi(n-m)t)}{\pi(n-m)} - i \frac{\cos(\pi(n-m)t)}{\pi(n-m)} \right]_{-1}^{1} = 0$$

<u>Exercise</u>: compute $\langle f_n, f_n \rangle_2$.

Thus for any $N \in \mathbb{N}$, $f_{-N}, f_{-(N-1)}, ..., f_{-1}, f_0, f_1, ..., f_N$ is an orthogonal system, and so by

the lemma it is linearly independent. Since we can increase N to however large we want, it follows that $C[-1,1]$ is _infinite dimensional_.

Fact "Fourier Analysis" says $\{f_n : n \in \mathbb{Z}\}$ is generating for $C[-1,1]$ (in a certain sense). □

Orthonormal Bases

Def An orthogonal (resp. orthonormal) basis is an orthogonal (resp. orthonormal) system that is also a basis (i.e. generating).

· Recall that given an arbitrary basis B for V, computing coordinate vectors can be hard. E.g. for $\vec{x} \in \mathbb{F}^n$

$$[\vec{x}]_B = [I]_S^B [\vec{x}]_S = ([I]_B^S)^{-1} \vec{x}$$

What makes orthogonal basis nice is that these are much easier to find.

Prop $^{5.10}$ Let V be an inner product space, and let $B = \{\vec{v}_1, \dots, \vec{v}_n\}$ be an orthogonal basis. Then for any $\vec{x} \in V$,

$$[\vec{x}]_B = (\alpha_1, \dots, \alpha_n)^T$$

where

$$* \qquad \alpha_k = \frac{\langle \vec{x}, \vec{v}_k \rangle}{\|\vec{v}_k\|^2} \qquad k = 1, \dots, n.$$

Proof Since B is a basis we know

$$\vec{x} = \alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n$$

for some scalars $\alpha_1, \dots, \alpha_n$. We must show these scalars are as in $(*)$. For $1 \le k \le n$, we have

$$\langle \vec{x}, \vec{v}_k \rangle = \langle \sum_{j=1}^{n} \alpha_j \vec{v}_j, \vec{v}_k \rangle = \sum_{j=1}^{n} \alpha_j \langle \vec{v}_j, \vec{v}_k \rangle = \alpha_k \langle \vec{v}_k, \vec{v}_k \rangle = \alpha_k \|\vec{v}_k\|^2.$$

Solving for α_k yields $(*)$. □

· Note that this proposition implies

$$\vec{x} = \sum_{k=1}^{n} \frac{\langle \vec{x}, \vec{v}_k \rangle}{\|\vec{v}_k\|^2} \vec{v}_k$$

for all $\vec{x} \in V$. If we further assume that $\vec{v}_1, \dots, \vec{v}_n$ is orthonormal, then this simplifies to

$$\vec{x} = \sum_{k=1}^{n} \langle \vec{x}, \vec{v}_k \rangle \vec{v}_k.$$

5.3 Orthogonal projections and Gram-Schmidt orthogonalization

Def Let V be an inner product space and let $E \subseteq V$ be a subspace. An orthogonal projection onto E is a linear transformation $P_E : V \to V$ satisfying

① $P_E(\vec{v}) \in E$ for all $\vec{v} \in V$.

② $(\vec{v} - P_E(\vec{v})) \perp E$ for all $\vec{v} \in V$.

EX Let $E = \operatorname{span}\{\vec{e}_2, \vec{e}_3\} \subseteq \mathbb{F}^3$. Then $P_E\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ y \\ z \end{pmatrix}$. Indeed, it clearly satisfies ① and

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} - P_E\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 0 \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} = x\vec{e}_1 \perp E.$$

Prop 5.11 Let V be an inner product space and let E be a subspace. Suppose $\vec{v}_1, \dots, \vec{v}_r$ is an orthogonal basis for E. Then the orthogonal projection onto E is given by

$$P_E(\vec{v}) = \sum_{k=1}^{r} \frac{\langle \vec{v}, \vec{v}_k \rangle}{\|\vec{v}_k\|^2} \vec{v}_k.$$

Proof First note that P_E is linear by the linearity of the inner product. Also,

$$P_E(\vec{v}) \in \operatorname{span}\{\vec{v}_1, \dots, \vec{v}_r\} = E$$

so P_E satisfies ①. So it remains to check ②. By the first Lemma in Section 5.2, it suffices to show $(\vec{v} - P_E(\vec{v})) \perp \vec{v}_k$ for each $k = 1, \dots, n$. We have

$$\langle \vec{v} - P_E(\vec{v}), \vec{v}_k \rangle = \langle \vec{v}, \vec{v}_k \rangle - \langle P_E(\vec{v}), \vec{v}_k \rangle$$

$$= \langle \vec{v}, \vec{v}_k \rangle - \left\langle \sum_{j=1}^{r} \frac{\langle \vec{v}, \vec{v}_j \rangle}{\|\vec{v}_j\|^2}, \vec{v}_k \right\rangle$$

$$= \langle \vec{v}, \vec{v}_k \rangle - \sum_{j=1}^{n} \frac{\langle \vec{v}, \vec{v}_j \rangle}{\|\vec{v}_j\|^2} \langle \vec{v}_j, \vec{v}_k \rangle$$

$$= \langle \vec{v}, \vec{v}_k \rangle - \frac{\langle \vec{v}, \vec{v}_k \rangle}{\|\vec{v}_k\|^2} \langle \vec{v}_k, \vec{v}_k \rangle$$

$$= \langle \vec{v}, \vec{v}_k \rangle - \langle \vec{v}, \vec{v}_k \rangle = 0.$$

EX Recall that in $\mathbb{C}^n$, $\langle \vec{x}, \vec{y} \rangle = \vec{y}^* \vec{x}$. Thus for a subspace $E \subseteq \mathbb{C}^n$ with orthogonal basis $\vec{v}_1, \dots, \vec{v}_r$ The Proposition says:

$$P_E(\vec{x}) = \sum_{k=1}^{r} \frac{1}{\|\vec{v}_k\|^2} (\vec{v}_k^* \vec{x}) \vec{v}_k = \sum_{k=1}^{r} \frac{1}{\|\vec{v}_k\|^2} \vec{v}_k (\vec{v}_k^* \vec{x})$$

Observe that $\vec{v}_k \vec{v}_k^* \in M_{n \times 1} \cdot M_{1 \times n} \subseteq M_{n \times n}$. It follows that

$$[P_E] = \sum_{k=1}^{r} \frac{1}{\|\vec{v}_k\|^2} \vec{v}_k \vec{v}_k^* \in M_{n \times n}.$$

Thm 5.12 Let V be an inner product space and let P_E be an orthogonal projection onto a subspace E. Then for all $\vec{v} \in V$ and all $\vec{x} \in E$

$$\|\vec{v} - P_E(\vec{v})\| \le \|\vec{v} - \vec{x}\|$$

That is, $P_E(\vec{v})$ minimizes the distance from $\vec{v}$ to E. Moreover, if for some $\vec{x} \in E$

$$\| \vec{v} - P_E(\vec{v}) \| = \| \vec{v} - \vec{x} \|$$

Then $P_E(\vec{v}) = \vec{x}$. Consequently, any orthogonal projection onto E is unique.

Proof For $\vec{v} \in V$ and $\vec{x} \in E$, note that

$$\vec{v} - \vec{x} = \underbrace{\vec{v} - P_E(\vec{v})}_{\perp E} + \underbrace{P_E(\vec{v}) - \vec{x}}_{\in E}$$

So by the Pythagorean theorem

$$\| \vec{v} - \vec{x} \|^2 = \| (\vec{v} - P_E(\vec{v})) + (P_E(\vec{v}) - \vec{x}) \|^2$$
$$= \| \vec{v} - P_E(\vec{v}) \|^2 + \| P_E(\vec{v}) - \vec{x} \|^2 \geq \| \vec{v} - P_E(\vec{v}) \|^2$$

Which proves the first part. Note that $\| \vec{v} - \vec{x} \| = \| \vec{v} - P_E(\vec{v}) \|$ is only possible if

$$\| P_E(\vec{v}) - \vec{x} \| = 0 \iff P_E(\vec{v}) - \vec{x} = \vec{0} \iff P_E(\vec{v}) = \vec{x}.$$

Finally, let P_E' be another orthogonal projection onto E. Then by the first part of the theorem

$$\| \vec{v} - P_E(\vec{v}) \| \leq \| \vec{v} - P_E'(\vec{v}) \|$$

and reversing the roles of P_E and P_E' we get $\| \vec{v} - P_E(\vec{v}) \| = \| \vec{v} - P_E'(\vec{v}) \|$. So by the second part of the theorem we have $P_E(\vec{v}) = P_E'(\vec{v})$. Since $\vec{v} \in V$ was arbitrary, this implies $P_E = P_E'$. □

Gram-Schmidt Orthogonalization Algorithm

- From the first proposition in this section, we see that we can always define P_E provided E has an orthogonal basis. In this subsection we show this is always the case.

The following algorithm takes in a linearly independent system $\vec{x}_1, \ldots, \vec{x}_n$ and captures an orthogonal system $\vec{v}_1, \ldots, \vec{v}_n$ s.t.

$$\text{span}\{\vec{x}_1, \ldots, \vec{x}_n\} = \text{span}\{\vec{v}_1, \ldots, \vec{v}_n\}$$

Moreover, for each $r \leq n$ it satisfies

$$\text{span}\{\vec{x}_1, \ldots, \vec{x}_r\} = \text{span}\{\vec{v}_1, \ldots, \vec{v}_r\}$$

Step 1 Set $\vec{v}_1 := \vec{x}_1$ and denote $E_1 := \text{span}\{\vec{x}_1\} = \text{span}\{\vec{v}_1\}$

Step 2 Set

$$\vec{v}_2 := \vec{x}_2 - P_{E_1}\vec{x}_2 = \vec{x}_2 - \frac{\langle \vec{x}_2, \vec{v}_1 \rangle}{\| \vec{v}_1 \|^2} \vec{v}_1$$

Note that $\vec{v}_2 \perp E_1$ by definition of P_{E_1}. In particular, $\vec{v}_2 \perp \vec{v}_1$. Denote $E_2 := \text{span}\{\vec{v}_1, \vec{v}_2\} = \text{span}\{\vec{x}_1, \vec{x}_2\}$. Note $\vec{x}_2 \notin \text{span}\{\vec{v}_1\}$ since $\vec{x}_1, \vec{x}_2$ are lin. indep. Hence $\vec{v}_2 \neq \vec{0}$.

Step 3 Set

$$\vec{v}_3 := \vec{x}_3 - P_{E_2}\vec{x}_3 = \vec{x}_3 - \frac{\langle \vec{x}_3, \vec{v}_1 \rangle}{\| \vec{v}_1 \|^2} \vec{v}_1 - \frac{\langle \vec{x}_3, \vec{v}_2 \rangle}{\| \vec{v}_2 \|^2} \vec{v}_2$$

Then $\vec{v}_3 \perp E_2$ and in particular $\vec{v}_3 \perp \vec{v}_1$, $\vec{v}_3 \perp \vec{v}_2$. Note that $\vec{x}_3 \notin E_2$ since $E_2 = \text{span}\{\vec{x}_1, \vec{x}_2\}$ and $\vec{x}_1, \vec{x}_2, \vec{x}_3$ are lin. indep. Thus $\vec{v}_3 \neq \vec{0}$.

$\vdots$

Step r+1 Suppose we have already iterated the above process r times to construct an orthogonal system $\vec{v}_1, \ldots, \vec{v}_r$ (of non-zero vectors) such that
$$E_r := \text{span}\{\vec{v}_1, \ldots, \vec{v}_r\} = \text{span}\{\vec{x}_1, \ldots, \vec{x}_r\}$$
Define
$$\vec{v}_{r+1} := \vec{x}_{r+1} - P_{E_r}\vec{x}_{r+1} = \vec{x}_{r+1} - \sum_{k=1}^{r} \frac{\langle \vec{x}_{r+1}, \vec{v}_k \rangle}{\|\vec{v}_k\|^2} \vec{v}_k$$

Then $\vec{v}_{r+1} \perp E_r$ and therefore $\vec{v}_{r+1} \perp \vec{v}_k$ for $k = 1, \ldots, r$. Also, since $\vec{x}_{r+1} \notin E_r$ by lin. indep. of $\vec{x}_1, \ldots, \vec{x}_r, \vec{x}_{r+1}$ we have $\vec{v}_{r+1} \neq \vec{0}$.

Iterating this process n times yields the desired orthogonal system $\vec{v}_1, \ldots, \vec{v}_n$. □

Ex In $\mathbb{R}^3$
$$\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \vec{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \qquad \vec{x}_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

Step 1 $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Step 2 $\vec{v}_2 = \vec{x}_2 - \frac{\langle \vec{x}_2, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} - \frac{0+1+2}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

Step 3 $\vec{v}_3 = \vec{x}_3 - \frac{\langle \vec{x}_3, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 - \frac{\langle \vec{x}_3, \vec{v}_2 \rangle}{\|\vec{v}_2\|^2} \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \frac{1+0+1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{-1+0+2}{2} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

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

So
$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \vec{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \qquad \vec{v}_3 = \begin{pmatrix} 1/2 \\ -1 \\ 1/2 \end{pmatrix}$$

Let $E := \text{span}\{\vec{x}_1, \vec{x}_2\} = \text{span}\{\vec{v}_1, \vec{v}_2\}$. Then $\vec{v}_1, \vec{v}_2$ is an orthogonal basis for E and so we know the formula for P_E. Lets write out its matrix representation (with respect to the standard basis S):
$$[P_E]_S^S = \frac{1}{\|\vec{v}_1\|^2} \vec{v}_1 \cdot \vec{v}_1^* + \frac{1}{\|\vec{v}_2\|^2} \vec{v}_2 \cdot \vec{v}_2^* = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot (1\ 1\ 1) + \frac{1}{2} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \cdot (-1\ 0\ 1)$$

$$= \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 5/6 & 1/3 & -1/6 \\ 1/3 & 1/3 & 1/3 \\ -1/6 & 1/3 & 5/6 \end{pmatrix}$$ □

Rem Note that if $\vec{u}_k := \frac{\vec{v}_k}{\|\vec{v}_k\|}$ for $k = 1, \ldots, n$ then $\vec{u}_1, \ldots, \vec{u}_n$ is an ortho**normal** system which satisfies
$$\text{span}\{\vec{u}_1, \ldots, \vec{u}_r\} = \text{span}\{\vec{x}_1, \ldots, \vec{x}_r\} \qquad r = 1, \ldots, n.$$ □

Orthogonal Complement

Def For a subset $S \subseteq V$, its <u>orthogonal complement</u> is the set
$$S^\perp := \{ \vec{x} \in V : \vec{x} \perp S \}$$

Prop 5.13 $S^\perp$ is always a subspace (even if S is not). Also
$$(S^\perp)^\perp = \text{span } S$$
In particular, $S = (S^\perp)^\perp$ if and only if S is a subspace.

~~Proof~~ Homework 10 □

Lemma 5.14 Let $E \subseteq V$ be a subspace. Then $E \cap E^\perp = \{\vec{0}\}$.

~~Proof~~ Let $\vec{x} \in E \cap E^\perp$ then
$$\underset{\in E}{\underbrace{\langle \vec{x}}, \overset{\in E^\perp}{\overbrace{\vec{x} \rangle}}} = 0$$
so $\vec{x} = \vec{0}$ by non-degeneracy. □

Thm 5.15 Let V be an inner product space and let $E \subseteq V$ be a subspace. Then for every $\vec{v} \in V$ there exists unique $\vec{v}_1 \in E$ and $\vec{v}_2 \in E^\perp$ s.t.
$$\vec{v} = \vec{v}_1 + \vec{v}_2.$$

Proof Let P_E be the orthogonal projection onto E. Given $\vec{v} \in V$, set $\vec{v}_1 := P_E(\vec{v})$ and $\vec{v}_2 := \vec{v} - P_E(\vec{v})$. Then $\vec{v}_1 \in E$ and $\vec{v}_2 \in E^\perp$ by def. of P_E. Also
$$\vec{v}_1 + \vec{v}_2 = P_E(\vec{v}) + \vec{v} - P_E(\vec{v}) = \vec{v}.$$
Suppose $\vec{w}_1 \in E$ and $\vec{w}_2 \in E^\perp$ also satisfy $\vec{w}_1 + \vec{w}_2 = \vec{v}$. Then
$$\vec{v}_1 + \vec{v}_2 = \vec{w}_1 + \vec{w}_2$$
$$\vec{v}_1 - \vec{w}_1 = \vec{w}_2 - \vec{v}_2$$
Note that $\vec{v}_1 - \vec{w}_1 \in E$ while $\vec{w}_2 - \vec{v}_2 \in E^\perp$. Thus $\vec{v}_1 - \vec{w}_1 = \vec{w}_2 - \vec{v}_2 \in E \cap E^\perp = \{\vec{0}\}$ by the Lemma.
So $\vec{v}_1 - \vec{w}_1 = \vec{0} \Rightarrow \vec{v}_1 = \vec{w}_1$ and $\vec{w}_2 - \vec{v}_2 = \vec{0} \Rightarrow \vec{w}_2 = \vec{v}_2$. Hence $\vec{v}_1$ and $\vec{v}_2$ are unique. □

- Whenever $E_1, E_2 \subseteq V$ are subspaces s.t. every $\vec{v} \in V$ can be written uniquely as $\vec{v} = \vec{v}_1 + \vec{v}_2$ for $\vec{v}_1 \in E_1$ and $\vec{v}_2 \in E_2$, we write
$$V = E_1 \oplus E_2.$$
  The above theorem says
$$V = E \oplus E^\perp.$$

Cor 5.16 For a subspace $E \leq V$,
$$\dim(E) + \dim(E^\perp) = \dim(V)$$

Proof Using Gram-Schmidt, we can find orthogonal bases B and C for E and $E^\perp$ respectively. Then $B \cup C$ is orthogonal and therefore lin. indep. The theorem tells us $B \cup C$ is generating. Hence $B \cup C$ is a basis for V. □