**Warmup.** **Lecture 22:** more geometry, orthogonalization

1. Recall that the standard inner product on $V = \mathbb{R}^n$ is dot product. Given two vectors $\mathbf{u}, \mathbf{v} \in V$, with $\mathbf{v} \neq 0$, the projection of $\mathbf{u}$ onto $\mathbf{v}$ is

$$\mathrm{proj}_{\mathbf{v}}(\mathbf{u}) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v}.$$

   (a) Let $\mathbf{u} = (1, 0, -2)$. For each of the following $\mathbf{v}$, compute $\mathrm{proj}_{\mathbf{v}}(\mathbf{u})$.

   (i) $\mathbf{v} = (3, -1, 4)$; (ii) $\mathbf{v} = (1, 0, 0)$; (iii) $\mathbf{v} = (0, 1, 0)$;

   (iv) $\mathbf{v} = (1, 1, 1)$; and (v) $\mathbf{v} = (-2, 0, 4)$.

   For (ii), (iii), and (v), can you make sense of your answers geometrically?

   (b) Under what circumstances is (i) $\mathrm{proj}_{\mathbf{v}}(\mathbf{u}) = \mathbf{u}$? (ii) $\mathrm{proj}_{\mathbf{v}}(\mathbf{u}) = \mathrm{proj}_{\mathbf{u}}(\mathbf{v})$?

   (c) A particle is traveling along the line $y = 2x$ in $\mathbb{R}^2$ such that its position at time $t$ is $\mathbf{p}(t) = (t, 2t)$.

   (a) At time $t$, what is the point on the $x$-axis that the particle is closest to? How far away is the particle from that point?

   (b) At time $t$, what is the point on the line $y = 5x$ that the particle is closest to? How far away is the particle from that point?

   [Hint: project $\mathbf{p}(t)$ onto the line by taking *any* vector $\mathbf{v}$ (besides $\mathbf{0}$) on the line, and computing $\mathrm{proj}_{\mathbf{v}}(\mathbf{p}(t))$.]

2. *Time permitting:* Define the standard inner product on $M_{k,\ell}(F)$ (where $F = \mathbb{R}$ or $\mathbb{C}$) by $\langle X, Y \rangle = \mathrm{tr}(\overline{Y}^t X)$, where $\overline{Y}$ means take the complex conjugate of the entries of $Y$. Check that this *is* an inner product.

[See next slide for examples]

**Examples for warmup 2.**
In $M_{2,3}(F)$,

$$\begin{aligned}
\left\langle \begin{pmatrix} 1 & 3 & 0 \\ i & 0 & 1+i \end{pmatrix}, \begin{pmatrix} 0 & 5+2i & 1 \\ 2-i & 6 & 4i \end{pmatrix} \right\rangle &= \operatorname{tr}\left( \overline{\begin{pmatrix} 0 & 5+2i & 1 \\ 2-i & 6 & 4i \end{pmatrix}}^t \begin{pmatrix} 1 & 3 & 0 \\ i & 0 & 1+i \end{pmatrix} \right) \\
&= \operatorname{tr}\left( \begin{pmatrix} \overline{0} & \overline{2-i} \\ \overline{5+2i} & \overline{6} \\ \overline{1} & \overline{4i} \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ i & 0 & 1+i \end{pmatrix} \right) \\
&= \operatorname{tr}\left( \begin{pmatrix} 0 & 2+i \\ 5-2i & 6 \\ 1 & -4i \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ i & 0 & 1+i \end{pmatrix} \right) \\
&= \operatorname{tr}\left( \begin{pmatrix} -1+2i & 0 & 1+3i \\ 5+4i & 15-6i & 6+6i \\ 5 & 3 & 4-4i \end{pmatrix} \right) \\
&= -1+2i+15-6i+4-4i = \boxed{18-8i}.
\end{aligned}$$

Recall that $E_{i,j}$ denotes the matrix with a 1 in row $i$ and column $j$ (and 0's elsewhere). Since $\overline{1}=1$ and $\overline{0}=0$, we have $\overline{E_{i,j}} = E_{i,j}$. So for any $k \geq 1$ and $\ell \geq 3$ (i.e. $M_{k,\ell}(F)$ is big enough to hold $E_{1,2}$ and $E_{1,3}$), we have

$$\langle E_{1,2}, E_{1,3}\rangle = \operatorname{tr}(\overline{E_{1,3}}^t E_{1,2}) = \operatorname{tr}(E_{3,1}E_{1,2}) = \operatorname{tr}(E_{3,2}) = \boxed{0}; \text{ and}$$
$$\langle E_{1,3}, E_{1,3}\rangle = \operatorname{tr}(\overline{E_{1,3}}^t E_{1,3}) = \operatorname{tr}(E_{3,1}E_{1,3}) = \operatorname{tr}(E_{3,3}) = \boxed{1}.$$

**Last time:** Let $V$ be a vector space over a field $F$ where $F=\mathbb{R}$ or $\mathbb{C}$.

An **inner product** on $V$ is a function

$$\begin{aligned} \langle\ ,\ \rangle \colon V \times V &\to F \\ (\mathbf{u},\mathbf{v}) &\mapsto \langle \mathbf{u},\mathbf{v}\rangle \end{aligned}$$

that is linear in the first coordinate, conjugate symmetric, and positive-definite. We can additionally prove that $\langle , \rangle$ is conjugate linear in the second coordinate, is non-degenerate, and has $\langle \mathbf{u}, 0\rangle = 0$ for all $\mathbf{u}$. We call $(V, \langle , \rangle)$ an **inner product space (IPS)**.

Favorite examples: The standard inner product on. . .

. . . $\mathbb{R}^n$ is dot product: $\langle \mathbf{u},\mathbf{v}\rangle = \mathbf{u}\cdot\mathbf{v}$;

. . . $\mathbb{C}^n$ is conjugate dot product: $\langle \mathbf{u},\mathbf{v}\rangle = \mathbf{u}\cdot\overline{\mathbf{v}}$;

. . . { continuous functions $f : [0,1] \to \mathbb{R}$ } is $\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt$.

We can use $\langle , \rangle$ to start building geometry on an inner product space.
First, we define the **norm** of $\mathbf{u} \in V$ by $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u},\mathbf{u}\rangle}$. Then $\|\cdot\|$ is also positive-definite, homogeneous, and satisfies the Pythagorean Theorem (where we define $\mathbf{u} \perp \mathbf{v}$ by $\langle \mathbf{u},\mathbf{v}\rangle = 0$) and the triangle inequality.

**Lemma.** For any $\mathbf{v} \neq \mathbf{0}$, $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} \overset{\text{def}}{=} \left(\frac{1}{\|\mathbf{v}\|}\right)\mathbf{v}$ is a **unit vector** (has length 1).

## Distance and metric spaces

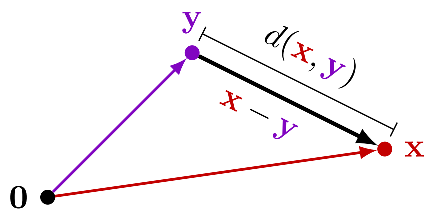
Still: Let $(V, \langle , \rangle)$ be an IPS/$F = \mathbb{R}$ or $\mathbb{C}$.

**Remember!!** We hold in our mind two notions about spaces like $\mathbb{R}^n$: an element is both a **point** *and* a **vector** pointing from $\mathbf{0}$ to that point.

---

Define the distance between points $\mathbf{x}, \mathbf{y} \in V$ as

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$




**Proposition.** For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$,

1. Symmetry: $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$.
2. Positive-definiteness: $d(\mathbf{x}, \mathbf{y}) \geq 0$, and $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$.
3. Triangle inequality: $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$.

Namely, an inner product space is also a **metric space**.

*Pf.* (exercise)

## Angles

Let $(V, \langle , \rangle)$ be an inner product space over $F = \mathbb{R}$

Let $\mathbf{x}, \mathbf{y}$ be nonzero vectors in $V$. Define the **angle** $\theta$ between $\mathbf{x}$ and $\mathbf{y}$ is defined to by

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\|\mathbf{y}\| \cos(\theta),$$

i.e.

$$\theta \stackrel{\text{def}}{=} \cos^{-1}\left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}\right).$$

**Wait!** Is this well-defined? Is this a reasonable definition?

Cauchy-Schwarz (last time) says $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$. So

$$-\|\mathbf{x}\|\|\mathbf{y}\| \leq \langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{x}\|\|\mathbf{y}\|,$$

and hence $-1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1$ (i.e. $\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}$ is in cosine's range). 

*Remark:*

$$\cos(\theta) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} = \left\langle \frac{1}{\|\mathbf{x}\|}\mathbf{x}, \overline{\left(\frac{1}{\|\mathbf{y}\|}\right)}\mathbf{y} \right\rangle = \left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{y}}{\|\mathbf{y}\|} \right\rangle,$$

because $\|\mathbf{y}\| \in \mathbb{R}$.
In particular, $\theta$ is a statistic in terms of the unit vectors $\frac{\mathbf{x}}{\|\mathbf{x}\|}$ and $\frac{\mathbf{y}}{\|\mathbf{y}\|}$.

## Orthogonal sets

Still: $(V, \langle\ ,\ \rangle)$ an IPS/$F = \mathbb{R}$ or $\mathbb{C}$.

A subset $S \subseteq V$ is **orthogonal** if its elements are pairwise orthogonal: if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ *for all* $\mathbf{u}, \mathbf{v} \in S$ with $\mathbf{u} \neq \mathbf{v}$.

If $S$ is an orthogonal subset of $V$ *and* $\|\mathbf{u}\| = 1$ for all $u \in S$, we say $S$ is an **orthonormal** subset of $V$. Note that orthonormal means

$$\langle \mathbf{u}, \mathbf{v} \rangle = \delta_{\mathbf{u},\mathbf{v}} \quad \text{for all } \mathbf{u}, \mathbf{v} \in S.$$

**Examples.**

- The standard basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of $F^n$ is orthonormal with respect to the standard inner product.
- $S = \left\{ \frac{1}{\sqrt{2}}(1,1), \frac{1}{\sqrt{2}}(1,-1) \right\}$ is orthonormal with respect to the standard inner product on $\mathbb{R}^2$.
- $S = \{\cos(2\pi x), \sin(2\pi x)\}$ is orthogonal in the space of continuous functions $\mathcal{C}(\mathbb{R})$ with respect to the inner product

$$\langle f, g \rangle := \int_0^1 f(t)g(t)\, dt,$$

  but is not ortho*normal* since $\int_0^1 \cos^2(2\pi t)\ dt = 1/2 \neq 1$.

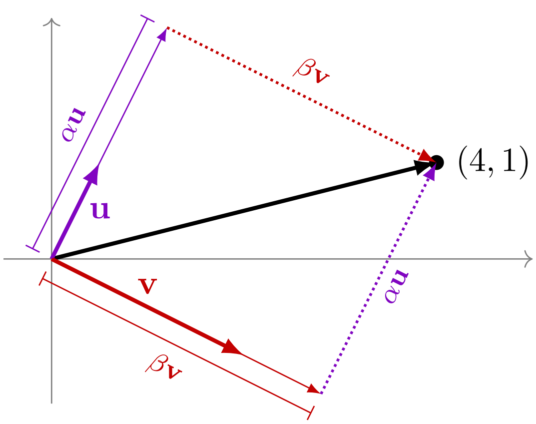
## Computing coordinates

**Question.** Given an orthonormal basis $\mathcal{B}$ of $V$, write the expansion of $\mathbf{v} \in V$ in terms of $\mathcal{B}$; i.e. what is $\mathrm{Rep}_{\mathcal{B}}(\mathbf{v})$?

**NOTATION:** Today, I'm going to use $\langle\!\langle, \rangle\!\rangle$ for ordered bases to emphasize the difference between this notation and inner-product notation.

**Example.** Let $\mathcal{B} = \langle\!\langle \mathbf{u}, \mathbf{v} \rangle\!\rangle$, where $\mathbf{u} = (1/2, 1)$ and $\mathbf{v} = (2, -1)$.

What is $\mathrm{Rep}_{\mathcal{B}}((4,1))$?




**Goal:**
solve for $\alpha$ and $\beta$

Notice!

$$\alpha\mathbf{u} = \mathrm{proj}_{\mathbf{u}}((4,1)) = \frac{\langle (4,1), \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u} \quad \text{and} \quad \beta\mathbf{v} = \mathrm{proj}_{\mathbf{v}}((4,1)) = \frac{\langle (4,1), \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v}.$$

## Computing coordinates

**Proposition.** Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be an orthogonal set of nonzero vectors in $V$, and let $\mathbf{y} \in FS$. Then

$$\mathbf{y} = \sum_{j=1}^{k} \frac{\langle \mathbf{y}, \mathbf{v}_j \rangle}{\langle \mathbf{v}_j, \mathbf{v}_j \rangle} \mathbf{v}_j = \sum_{j=1}^{k} \frac{\langle \mathbf{y}, \mathbf{v}_j \rangle}{\|\mathbf{v}_j\|^2} \mathbf{v}_j.$$

*Conceptually:* each term is just the projection of $\mathbf{y}$ onto $\mathbf{v}_j$.

*Computationally:* Since $\mathbf{y} \in FS$, there are $c_1, \ldots, c_k$ such that

$$\mathbf{y} = \sum_{i=1}^{k} c_i \mathbf{v}_i.$$

Plug this into $\langle \mathbf{y}, \mathbf{v}_j \rangle$ and see what happens!

**Corollary 1.** If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is orthonormal and $\mathbf{y} \in FS$, then

$$\mathbf{y} = \sum_{j=1}^{k} \langle \mathbf{y}, \mathbf{v}_j \rangle \mathbf{v}_i.$$

*Pf.* $\langle \mathbf{v}_j, \mathbf{v}_j \rangle = 1$ for all $j$.

**Corollary 2.** If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthogonal set of nonzero vectors in $V$ then $S$ is linearly independent.

*Pf.* Apply the Prop. to $\mathbf{y} = \mathbf{v}_i$ and $S' = S - \{\mathbf{v}_i\}$ to see $\mathbf{v}_i \notin FS'$.

## Gram-Schmidt orthognalization

**Goal.** Build an orthonormal subset or basis from a previously existing linearly independent set. (Recall that we learned how to recursively build a linearly independent subset of any set earlier this semester—Lectures 5 & 6 or so.)

**Algorithm.**

INPUT: $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$, a linearly independent subset of $V$.
Start with

$$\mathbf{v}_1 := \mathbf{w}_1.$$

For $k = 2, 3, \ldots, n$, define $\mathbf{v}_k$ by starting with $\mathbf{w}_k$, then subtracting off the components of $\mathbf{w}_k$ along the previously found $\mathbf{v}_i$ :

$$\mathbf{v}_k := \mathbf{w}_k - \sum_{i=1}^{k-1} \frac{\langle \mathbf{w}_k, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i.$$

[Check: Is $\langle \mathbf{v}_k, \mathbf{v}_i \rangle = 0$ for all $i \leq k$?]

OUTPUT: $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthogonal set with $FS' = FS$.

or

OUTPUT: $S'' = \left\{ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \ldots, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|} \right\}$ an orthonormal set with $FS'' = FS$.

# Example

$V = \mathcal{P}_1(\mathbb{R})$ with inner product

$$\langle f, g \rangle = \int_0^1 f(t)g(t)\,dt.$$

Apply Gram-Schmidt to the basis $\langle\langle 1, x \rangle\rangle$ to get an orthonormal basis.
Start with $\mathbf{v}_1 = 1$, then let

$$\mathbf{v}_2 = x - \frac{\langle x, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = x - \frac{\langle x, 1 \rangle}{\|1\|^2} \cdot 1$$

$$= x - \frac{\int_0^1 t\,dt}{\int_0^1 dt} \cdot 1 = x - \frac{1}{2}.$$

Check orthogonality:

$$\langle 1, x - 1/2 \rangle = \int_0^1 (t - 1/2)\,dt = 0.$$

Orthogonal basis: $\langle\langle 1, x - \frac{1}{2} \rangle\rangle$. Scale to get an orthonormal basis:

$$\begin{aligned}
\|\mathbf{v}_1\|^2 &= \int_0^1 dt = 1 \\
\|\mathbf{v}_2\|^2 &= \langle x - 1/2, x - 1/2 \rangle \\
&= \int_0^1 (t - 1/2)^2\,dt \\
&= \left.\frac{(t-1/2)^3}{3}\right|_{t=0}^{1} = \frac{1/8}{3} - \frac{-1/8}{3} = \frac{1}{12}.
\end{aligned}$$

Orthonormal basis: $\left\langle\!\left\langle 1, \sqrt{12}(x - 1/2) \right\rangle\!\right\rangle$

You try.

1. Consider the standard basis $\mathcal{B} = \langle\!\langle 1, x, x^2 \rangle\!\rangle$ of $\mathcal{P}_2(\mathbb{R})$.
   (a) Check (for yourself) that the first two steps of Gram-Schidt go the same as in the example we just did. Then finish the algorithm to get an orthonormal basis $\mathcal{B}' = \langle\!\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle\!\rangle$ of $\mathcal{P}_3(\mathbb{R})$.
   (b) Check your answer by computing $\langle \mathbf{v}_3, \mathbf{v}_1 \rangle$ and $\langle \mathbf{v}_3, \mathbf{v}_2 \rangle$. (You should get 0.)
   (c) Compute the associated orthonormal basis $\mathcal{B}''$ of $\mathcal{P}_3(\mathbb{R})$. (Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are the same as in the example above, you only need to compute $\|\mathbf{v}_3\|$.)
2. For $\mathcal{P}_n(\mathbb{R})$, we're finding that the standard basis is *not* orthonormal with respect to the standard inner product. What about our other favorite examples? For each of the following, either verify that the standard ordered basis *is* orthonormal (with respect to the standard inner product), or apply Gram-Schmidt orthogonalization to the standard ordered basis to get an orthonormal basis.
   (a) $V = \mathbb{R}^3$ (with $F = \mathbb{R}$ and $\langle , \rangle$ being dot product);
   (b) $V = \mathbb{C}^3$ (with $F = \mathbb{R}$ and $\langle , \rangle$ being conjugate dot product);
   (c) $V = M_{2,3}(\mathbb{R})$ (with $F = \mathbb{R}$ and $\langle , \rangle$ the inner product in the warmup);
   (d) $V = M_{2,3}(\mathbb{C})$ (with $F = \mathbb{C}$ and $\langle , \rangle$ the inner product in the warmup).