Math 2270-002 Week 12 notes
We will not necessarily finish the material from a given day's notes on that day. We may also add or subtract some material as the week progresses, but these notes represent an in-depth outline of what we plan to cover. These notes cover material in 6.1-6.4

Chapter 6 is about orthogonality and related topics. We'll spend maybe two weeks plus a day in this chapter. The ideas we develop start with the dot product, which we've been using algebraically to compute individual entries in matrix products, but which has important geometric meaning. By the end of the Chapter we will see applications to statistics, discuss generalizations of the dot product, "inner products", which can apply to function vector spaces and which lie at the heart of physics applications that use Fourier series, and more recent applications such as image and audio compression, see e.g.

https://en.wikipedia.org/wiki/Discrete_cosine_transform

Mon Nov 12

- 6.1-6.2 dot product, length, orthogonality, projection onto the span of a single vector, and angles - in $\mathbb{R}^n$.

Announcements:
- HW due Wed, but it's short.
- I'll try to grade the exams for Tuesday for sure by Wed.

Warm-up Exercise: no warmup, but you could look over the new HW assignment, due this Wed.

Recall, for any two vectors $\underline{\boldsymbol{v}}, \underline{\boldsymbol{w}} \in \mathbb{R}^n$, the dot product $\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}}$ is the scalar computed by the definition

$$\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}} := \sum_{i=1}^{n} v_i\, w_i\,.$$

We don't care if $\underline{\boldsymbol{v}}, \underline{\boldsymbol{w}}$ are row vectors or column vectors, or one of each, for the dot product.

We've been using the dot product algebraically to compute entries of matrix products $A\,B$, since

$$entry_{i\,j}\,[A\,B] = \left[row_i\,A\right]\left[col_j\,B\right] = \left[row_i\,A\right] \cdot \left[col_j\,B\right].$$

The algebra for dot products is a mostly a special case of what we already know for matrices, but worth writing down and double-checking, so we're ready to use it in the rest of Chapters 6 and 7.

Exercise 1 Check why

1a) dot product is commutative:

$$\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}} = \underline{\boldsymbol{w}} \cdot \underline{\boldsymbol{v}}\,.$$

1b) dot product distributes over addition:

$$(\underline{\boldsymbol{u}} + \underline{\boldsymbol{v}}) \cdot \underline{\boldsymbol{w}} = \underline{\boldsymbol{u}} \cdot \underline{\boldsymbol{w}} + \underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}} \quad \leftarrow \quad (\vec{u}+\vec{v})\cdot\vec{w} = \sum_{i=1}^{n} \underbrace{(u_i+v_i)w_i}_{u_i w_i + v_i w_i}$$

$$= \sum_{i=1}^{n} u_i w_i + \sum_{i=1}^{n} v_i w_i$$

$$= \vec{u}\cdot\vec{w} + \vec{v}\cdot\vec{w}$$

$$\underline{\boldsymbol{u}} \cdot (\underline{\boldsymbol{v}} + \underline{\boldsymbol{w}}) = \underline{\boldsymbol{u}} \cdot \underline{\boldsymbol{v}} + \underline{\boldsymbol{u}} \cdot \underline{\boldsymbol{w}}$$

1c) for $k \in \mathbb{R}$,

$$(k\,\underline{\boldsymbol{v}}) \cdot \underline{\boldsymbol{w}} = k\,(\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}}) = \underline{\boldsymbol{v}} \cdot (k\,\underline{\boldsymbol{w}}).$$

all three of these equal $k \sum_{i=1}^{n} v_i w_i$

1d) dot product distributes over linear combinations:

$$\underline{\boldsymbol{v}} \cdot \left(c_1\,\underline{\boldsymbol{w}}_1 + c_2\,\underline{\boldsymbol{w}}_2 + \ldots + c_k\,\underline{\boldsymbol{w}}_k\right) = c_1\left(\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}}_1\right) + c_2\left(\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}}_2\right) + \ldots + c_k\left(\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}}_k\right).$$

combine 1b), 1c)

$$\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}} := \sum_{i=1}^{n} v_i\, w_i$$

1e)

$$\sum_{i=1}^{n} v_i^2$$

$$\shortparallel$$

0 scalar

$$\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{v}} > 0 \quad \text{for each } \underline{\boldsymbol{v}} \neq \underline{\boldsymbol{0}} \quad (\text{and } \underline{\boldsymbol{0}} \cdot \underline{\boldsymbol{0}} = \cancel{\underline{\boldsymbol{0}}}\,)$$

Chapter 6 is about topics related to the geometry of the dot product. It begins now, with definitions and consequences that generalize what you learned for $\mathbb{R}^2$, $\mathbb{R}^3$ in your multivariable Calculus class, to $\mathbb{R}^n$.

2) Geometry of the dot product, stage 1. We'll add examples with pictures as we go throught these definitions.

2a) For $\underline{\boldsymbol{v}} \in \mathbb{R}^n$ we define the *length* or *norm* or *magnitude* of $\underline{\boldsymbol{v}}$ by

$$\|\underline{\boldsymbol{v}}\| := \sqrt{\sum_{i=1}^{n} v_i^2} = (\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{v}})^{\frac{1}{2}} .$$

$$\left\| \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \right\| = \sqrt{1+4+9+16} = \sqrt{30}$$

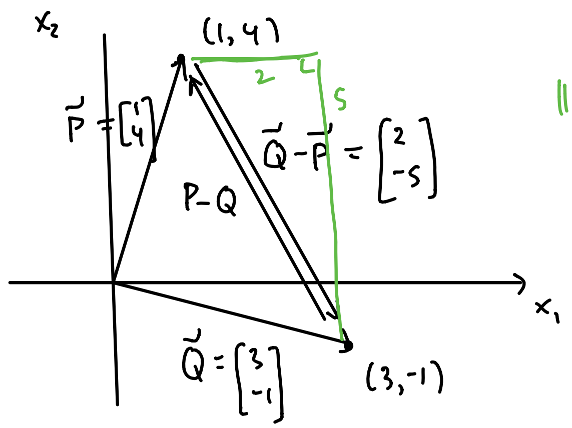
Notice that the length of a scalar multiple of a vector is what you'd expect:

$$\| t\,\underline{\boldsymbol{v}}\| = (t\,\underline{\boldsymbol{v}} \cdot t\,\underline{\boldsymbol{v}})^{\frac{1}{2}} = \left(t^2\, \underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{v}}\right)^{\frac{1}{2}} = |\,t\,|\,\|\underline{\boldsymbol{v}}\| .$$

~~Also notice that $\|\underline{\boldsymbol{v}}\| > 0$ unless $\underline{\boldsymbol{v}} = \underline{\boldsymbol{0}}$.~~

$$\left\| \begin{bmatrix} -5 \\ 25 \end{bmatrix} \right\| = \left\| 5 \begin{bmatrix} -1 \\ 5 \end{bmatrix} \right\| = 5\sqrt{26}$$

2b) The distance between points (with position vectors) $\underline{\boldsymbol{P}}$, $\underline{\boldsymbol{Q}}$ is defined to be the magnitude of the displacement vector(s) between them, $\|\underline{\boldsymbol{Q}} - \underline{\boldsymbol{P}}\|$ (or $\|\underline{\boldsymbol{P}} - \underline{\boldsymbol{Q}}\|$).

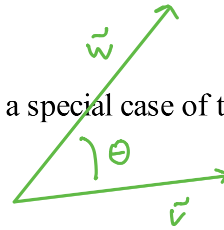


$$\|Q-P\| = \sqrt{29} = \sqrt{2^2+5^2}$$

2c) For $\underline{v}, \underline{w} \in \mathbb{R}^n$, we define $\underline{v}$ to be *orthogonal* (or *perpendicular*) to $\underline{w}$ if and only if

$$\underline{v} \cdot \underline{w} = 0.$$

And in this case we write $\underline{v} \perp \underline{w}$.



Note: In $\mathbb{R}^2$ or $\mathbb{R}^3$ and in your multivariable calculus class, this definition was a special case of the identity

$$\underline{v} \cdot \underline{w} = \|\underline{v}\| \ \|\underline{w}\| \cos(\theta)$$

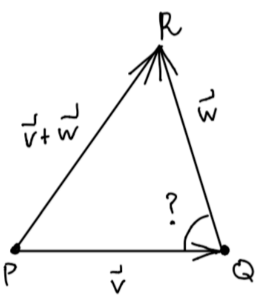
where θ is the angle between $\underline{v}, \underline{w}$. (Because $\cos(\theta) = 0$ when $\theta = \frac{\pi}{2}$.) That identity followed from the law of cosines, although you probably don't recall the details. Today we'll use the identity above to *define* angles between vectors, in $\mathbb{R}^n$, and show that it makes sense. (Next week, in section 6.6 we'll see that what we call "$\cos(\theta)$" in Math 2270 is known as the "correlation coefficient" in linear regression problems in statistics. In about two weeks, we will use the same identity to define angles between functions and perpendicular functions, in inner product function spaces, sections 6.8-6.9.)

$$\cos\theta = \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\| \cdot \|\vec{w}\|}$$

(= "r" the correlation coeff between $\{x_i\}_{i=1}^n$, $\{y_i\}_{i=1}^n$)

2d) The $\mathbb{R}^n$ reason for defining orthogonality as in 2c is that the Pythagorean Theorem holds for the triangle with displacement vectors $\underline{v}, \underline{w}$ and hypotenuse $\underline{v} + \underline{w}$ if and only if $\underline{v} \cdot \underline{w} = 0$. Check!

$$\begin{aligned}
\|\vec{v} + \vec{w}\|^2 &= (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) \\
&= (\vec{v} + \vec{w}) \cdot \vec{v} + (\vec{v} + \vec{w}) \cdot \vec{w} \\
&= \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{v} + \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{w} \\
\|\vec{v} + \vec{w}\|^2 &= \|\vec{v}\|^2 + 2\,\vec{v} \cdot \vec{w} + \|\vec{w}\|^2
\end{aligned}$$



Pythag. theorem if and only if $\vec{v} \cdot \vec{w} = 0$

$\pi/2 = ?$ $\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2$

2e) A vector $\underline{\boldsymbol{u}} \in \mathbb{R}^n$ is called a *unit vector* if and only if $\|\underline{\boldsymbol{u}}\| = 1$.

$\begin{bmatrix} 1/2 \\ \sqrt{3}/2 \end{bmatrix}$ is a unit vector $\|\cdot\|^2 = \frac{1}{4} + \frac{3}{4} = 1.$

$= \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ for $\theta = \pi/3$

2f) If $\underline{\boldsymbol{v}} \in \mathbb{R}^n$ then the unit vector in the direction of $\underline{\boldsymbol{v}}$ is given by

$$\underline{\boldsymbol{u}} = \frac{1}{\|\underline{\boldsymbol{v}}\|}\underline{\boldsymbol{v}}.$$

$\left\| \frac{\vec{v}}{\|\vec{v}\|} \right\| = \frac{1}{\|\vec{v}\|}\|\vec{v}\| = 1$

using scalar mult. prop

unit vector in dir. $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ is $\frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ -1 \end{bmatrix}$

2g) Projection onto a line. Let $\underline{\boldsymbol{v}} \in \mathbb{R}^n$ be a non-zero vector, let $L = span\{\underline{\boldsymbol{v}}\}$ be a line through the origin. Then for any $\underline{\boldsymbol{x}} \in \mathbb{R}^n$ the projection of $\underline{\boldsymbol{x}}$ onto L is defined by the formula

$$proj_L\, \underline{\boldsymbol{x}} := (\underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{u}})\, \underline{\boldsymbol{u}} \quad = \left(\vec{x} \cdot \frac{\vec{v}}{\|\vec{v}\|}\right) \frac{\vec{v}}{\|\vec{v}\|}$$

for $\underline{\boldsymbol{u}}$ the unit vector in the direction of $\underline{\boldsymbol{v}}$, $\underline{\boldsymbol{u}} = \frac{1}{\|\underline{\boldsymbol{v}}\|}\underline{\boldsymbol{v}}$. Equivalently

$$proj_L\, \underline{\boldsymbol{x}} := \frac{(\underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{v}})}{\|\underline{\boldsymbol{v}}\|^2}\underline{\boldsymbol{v}}.$$

Then the vector

$$\underline{\boldsymbol{z}} := \underline{\boldsymbol{x}} - (\underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{u}})\, \underline{\boldsymbol{u}}$$

is perpendicular to every vector in $span\{\underline{\boldsymbol{v}}\} = span\{\underline{\boldsymbol{u}}\}$. Thus every triangle containing (the point with position vector) $\underline{\boldsymbol{x}}$, (the point with position vector) $proj_L\, \underline{\boldsymbol{x}}$, and another point (with position vector) $\underline{\boldsymbol{w}}$ on the line L is a Pythagorean triangle. Consequently, $proj_L \underline{\boldsymbol{x}}$ is the (position vector of) nearest point on L to (the point with position vector) $\underline{\boldsymbol{x}}$.

$\vec{z} \perp L = span\{\vec{u}\}$:

iff $\vec{z} \cdot (t\vec{u}) = 0$ for all t

iff $\vec{z} \cdot \vec{u} = 0$

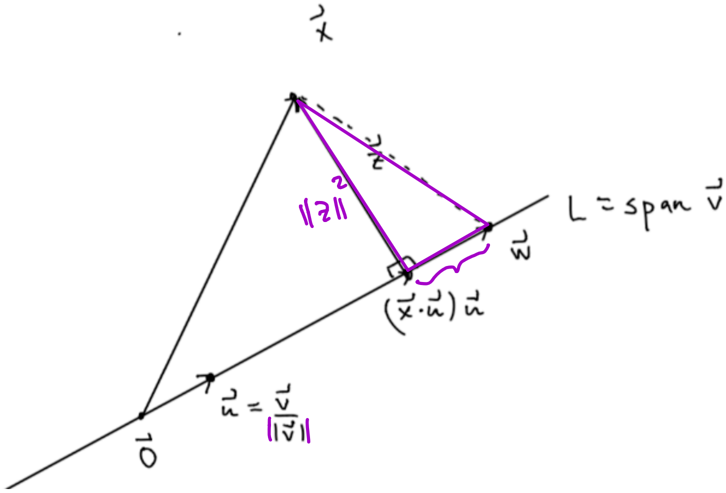
iff $(\vec{x} - (\vec{x}\cdot\vec{u})\vec{u}) \cdot \vec{u} \stackrel{?}{=} 0$

$\| $

$\vec{x}\cdot\vec{u} - (\vec{x}\cdot\vec{u})\underbrace{\vec{u}\cdot\vec{u}}_{1}$

$\vec{x}\cdot\vec{u} - \vec{x}\cdot\vec{u}$

$\stackrel{\checkmark}{=} 0$



So if $\vec{w} \in span\{\vec{u}\}$

then $\vec{w} - (\vec{x}\cdot\vec{u})\vec{u} = t\vec{u} \perp \vec{z}$

So $\|x - w\|^2 = \|\vec{z}\|^2 + t^2\|\vec{u}\|^2 > \|\vec{z}\|^2$ unless $t = 0$

2h) Refer to the same diagram as in 2g, which is an $\mathbb{R}^n$ picture. Using the Pythagorean triangle with edges $(\underline{\mathbf{x}} \cdot \underline{\mathbf{u}})\underline{\mathbf{u}}$, $\underline{z}$, $\underline{\mathbf{x}}$ we have

$$\| (\underline{\mathbf{x}} \cdot \underline{\mathbf{u}})\underline{\mathbf{u}} \|^2 + \|\underline{z}\|^2 = \|\underline{\mathbf{x}}\|^2, \text{ i.e.}$$

$$(\underline{\mathbf{x}} \cdot \underline{\mathbf{u}})^2 + \| \underline{z} \|^2 = \|\underline{\mathbf{x}}\|^2.$$

The quantity $\underline{\mathbf{x}} \cdot \underline{\mathbf{u}}$ is called *the component of* $\underline{\mathbf{x}}$ in the direction of $\underline{\mathbf{u}}$, and from the formula above,

$$-\|x\| \le \underline{\mathbf{x}} \cdot \underline{\mathbf{u}} \le \| x\| \ .$$

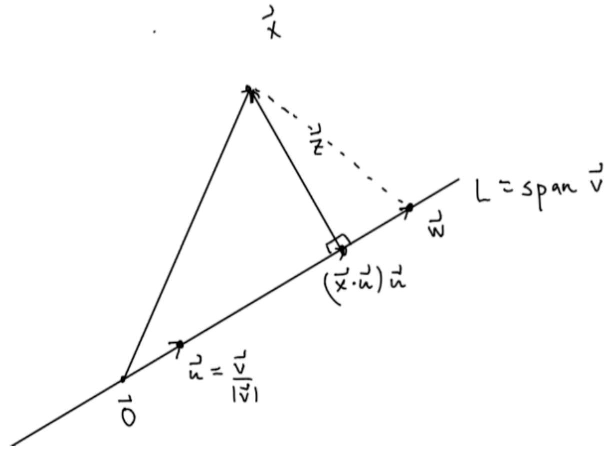
Define the angle θ between $\underline{\mathbf{v}}$ and $\underline{\mathbf{w}}$ the same way we would in $\mathbb{R}^2$, using the congruent triangle in the figure below, namely

$$\cos(\theta) = \frac{(\underline{\mathbf{x}} \cdot \underline{\mathbf{u}})}{\|\underline{\mathbf{x}}\|}.$$

Notice that $-1 \le \cos(\theta) \le 1$ and so there is a unique θ with $0 \le \theta \le \pi$ for which the cos θ equation can hold. Substituting $\underline{\mathbf{u}} = \dfrac{\underline{\mathbf{v}}}{\|\underline{\mathbf{v}}\|}$ gives the familiar formulas that you learned in multivariable Calculus for $\mathbb{R}^2$, $\mathbb{R}^3$, which now holds in $\mathbb{R}^n$.

$$\cos(\theta) = \frac{\left(\underline{\mathbf{x}} \bullet \dfrac{\underline{\mathbf{v}}}{\|\underline{\mathbf{v}}\|}\right)}{\|\underline{\mathbf{x}}\|} = \frac{(\underline{\mathbf{x}} \bullet \underline{\mathbf{v}})}{\|\underline{\mathbf{x}}\| \, \|\underline{\mathbf{v}}\|}, \text{ i.e.}$$

$$(\underline{\mathbf{x}} \bullet \underline{\mathbf{v}}) = \|\underline{\mathbf{x}}\| \, \|\underline{\mathbf{v}}\| \cos(\theta)$$



3) Summary exercise

a) In $\mathbb{R}^2$, let $L = span\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$. Find $proj_L\begin{bmatrix} 3 \\ 4 \end{bmatrix}$. Illustrate.

b) Verify the Pythagorean Theorem for some triple of points where two of them are $(3, 4)$, the point with position vector $proj_L\begin{bmatrix} 3 \\ 4 \end{bmatrix}$, and the third one is any other point on $L = span\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$.

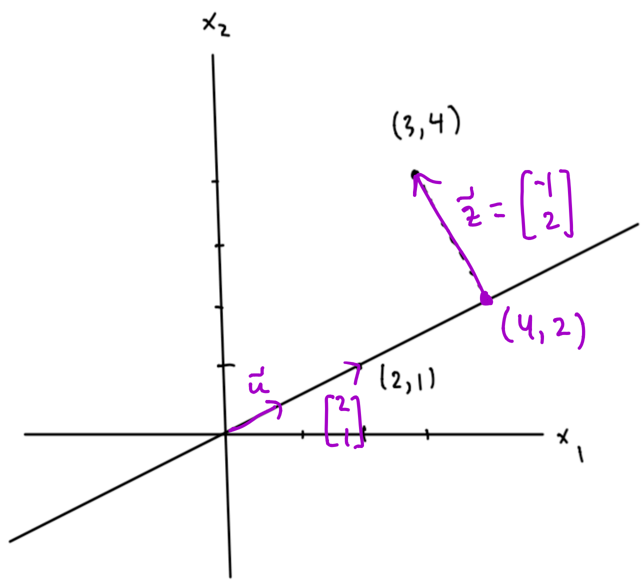
$proj_L\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ $\qquad$ $\vec{u}$ for the line : $\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$$= (\vec{x} \cdot \vec{u})\vec{u}$$

$$= \left(\begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right)\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= \frac{1}{5}\underbrace{(10)}_{\begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$



$\begin{bmatrix} -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 0$