HW for Friday
4.3
Chapter 4 Review
5.1
$\to$ but postpone the several 5.2 problems

Math 2270-3
Wed Oct 21.

Finish §5.1: Orthonormal bases and projections.

Recall from yesterday.

<u>Def.</u> $\{\vec{u}_1, \vec{u}_2, \ldots \vec{u}_k\} \subset \mathbb{R}^n$ is an <u>orthonormal</u> set iff $\vec{u}_i \cdot \vec{u}_j = \begin{cases} 1 & i=j \\ 0 & i \neq j. \end{cases}$

<u>Good Theorem</u>: Let $W = \operatorname{span}\{\vec{u}_1, \vec{u}_2, \ldots \vec{u}_k\}$, with $\{\vec{u}_1, \vec{u}_2, \ldots \vec{u}_k\}$ orthonormal. Then

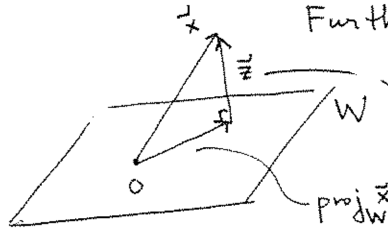
a) $\{\vec{u}_1, \vec{u}_2, \ldots \vec{u}_k\}$ are linearly ind., so basis for W. Call it $\mathcal{B}$.

b) If $\vec{v} \in W$, then $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} \vec{v}\cdot\vec{u}_1 \\ \vec{v}\cdot\vec{u}_2 \\ \vdots \\ \vec{v}\cdot\vec{u}_n \end{bmatrix}$

c) If $\vec{x} \in \mathbb{R}^n$, then $\operatorname{proj}_W \vec{x} := \sum_{j=1}^{k} (\vec{x}\cdot\vec{u}_j)\vec{u}_j$ is the nearest point in W to $\vec{x}$.

Furthermore, $\vec{z} := \vec{x} - \operatorname{proj}_W \vec{x}$ is $\perp$ to every vector in W.

We write $z \perp W$.



(sometimes write $\vec{x}^{\perp}$).

$\operatorname{proj}_W \vec{x}$ (sometimes write $\vec{x}^{\parallel}$)

$$\vec{x} = \operatorname{proj}_W \vec{x} + \vec{z} = \vec{x}^{\parallel} + \vec{x}^{\perp}$$
$$\|\vec{x}\|^2 = \|\vec{x}^{\parallel}\|^2 + \|\vec{x}^{\perp}\|^2$$

Do proof! p.5 Monday.

Then do Example 4 from Monday notes:

4a) Check $\mathcal{B} := \left\{ \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \frac{1}{3}\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \right\}$ is o.n. basis for $\mathbb{R}^3$

4b) for $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, find $[\vec{x}]_{\mathcal{B}}$ & check ans.

In Chapter 2, we considered the projection-onto-a-line-thru-the-origin linear transformation. ②

This generalizes! Fill in these details!

Theorem Let $\{\vec{u}_1, \vec{u}_2, \dots \vec{u}_k\}$ be an orthonormal basis for $W \subset \mathbb{R}^n$

Define $T : \mathbb{R}^n \to \mathbb{R}^n$

$$T(\vec{x}) := \text{proj}_W \vec{x} = \sum_{j=1}^{k} (\vec{u}_j \cdot \vec{x})\, \vec{u}_j$$

Then

a) T is a linear transformation

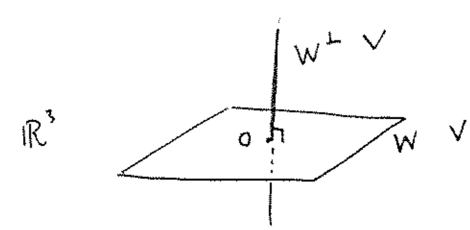
a)

b)

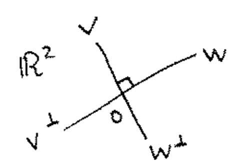
b) $\text{Image}(T) = W$

c) Define $W^\perp := \{\vec{x} \in \mathbb{R}^n \text{ s.t. } \vec{x} \cdot \vec{w} = 0 \ \forall \vec{w} \in W\}$ "W perp" "orthogonal complement of W"

Then $\ker(T) = W^\perp$



d) $\dim(W) + \dim(W^\perp) = n$



e) the intersection $W \cap W^\perp = \{\vec{0}\}$

proof: let $\vec{x} \in W \cap W^\perp$

Then $\underset{W}{\underset{\cap}{\vec{x}}} \cdot \underset{W^\perp}{\underset{\cap}{\vec{x}}} = 0$! ∎

f) $(W^\perp)^\perp = W$.

proof: $W \subset (W^\perp)^\perp$ since every $w \in W$ is $\perp$ to each $\vec{z} \in W^\perp$.

but $\dim(W^\perp) + \dim((W^\perp)^\perp) = n$, so $\dim(W) = \dim((W^\perp)^\perp)$

Thus $W = (W^\perp)^\perp$ ∎

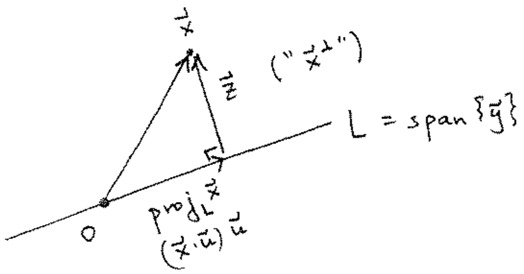
Notice, to find $W^\perp$, you don't need an o.n. basis for W.
In fact, you've had problems like this on previous HW

<u>Example 5</u> : Let $W = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \right\}$.

Find $W^\perp$

Cauchy-Schwarz inequality, triangle inequality, and angles <u>in $\mathbb{R}^n$</u> and why you care

Let $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\vec{x}, \vec{y} \neq \vec{0}$.



$$\vec{x} = \text{proj}_L \vec{x} + \vec{x}^\perp = \vec{x}^{/\!/} + \vec{x}^\perp$$

Since this is a right-triangle decomposition,

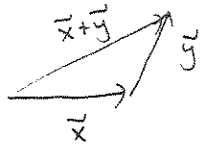
$\|\underbrace{\text{proj}_L \vec{x}}_{\left(\vec{x}\cdot\frac{\vec{y}}{\|\vec{y}\|}\right)\frac{\vec{y}}{\|\vec{y}\|}}\| \leq \|\vec{x}\|$ equality iff $\vec{x}^\perp = \vec{0}$, iff $\vec{x} /\!/ \vec{y}$.

$$\Rightarrow \frac{|\vec{x}\cdot\vec{y}|}{\|\vec{y}\|} \leq \|\vec{x}\|$$

$\boxed{|\vec{x}\cdot\vec{y}| \leq \|\vec{x}\|\,\|\vec{y}\|}$ equality iff $\vec{x} /\!/ \vec{y}$ (or $\vec{x}$ or $\vec{y} = \vec{0}$)

Cauchy Schwarz inequality

C.S. $\Rightarrow \Delta$ inequality

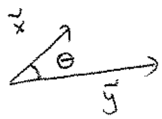


$$
\begin{aligned}
\|\vec{x}+\vec{y}\|^2 &= (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y}) \\
&= \vec{x}\cdot\vec{x} + 2\,\vec{x}\cdot\vec{y} + \vec{y}\cdot\vec{y} \\
\|\vec{x}+\vec{y}\|^2 &= \|\vec{x}\|^2 + 2\,\vec{x}\cdot\vec{y} + \|\vec{y}\|^2 \\
&\le \|\vec{x}\|^2 + 2\,|\vec{x}\cdot\vec{y}| + \|\vec{y}\|^2 \\
&\le \|\vec{x}\|^2 + 2\,\|\vec{x}\|\|\vec{y}\| + \|\vec{y}\|^2 \qquad \text{C.S.} \\
&= \left(\|\vec{x}\| + \|\vec{y}\|\right)^2
\end{aligned}
$$

$\sqrt{\ }$: $\boxed{\|\vec{x}+\vec{y}\| \le \|\vec{x}\| + \|\vec{y}\|}$ equality iff $\vec{x}$ & $\vec{y}$ positive scalar mults of each other (or one is zero).

triangle inequality.

angles in $\mathbb{R}^2$, $\mathbb{R}^3$, we used trig to show $\cos\theta = \dfrac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|}$



$0 \le \theta \le \pi$

in $\mathbb{R}^n$, we define the angle between $\vec{x}$, $\vec{y}$, $\measuredangle \vec{x},\vec{y}$ to be the unique θ, $0 \le \theta \le \pi$ for which

$$\cos\theta = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|}$$

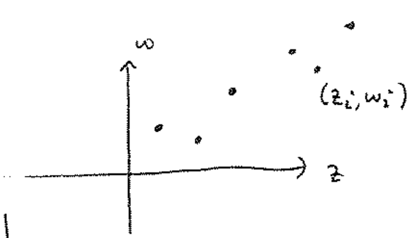
C.S. makes this possible, because $-1 \le \dfrac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|} \le 1$!

Example 6

Show $\measuredangle \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \pi/3$, by computing $\cos\theta$

Example 7 : Correlation coefficients "r" are angle cosines in Statistics!

Suppose you have n data points, $(z_1,w_1), (z_2,w_2), \ldots (z_n,w_n)$ with average $(\bar{z},\bar{w}) = \left(\frac{1}{n}\sum z_i, \frac{1}{n}\sum w_i\right)$



Then the regression coeff r is defined by

$$r = \frac{\sum_{i=1}^{n}(z_i-\bar{z})(w_i-\bar{w})}{\left(\sum z_i^2\right)^{1/2}\left(\sum w_i^2\right)^{1/2}} = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|} = \cos\theta \,!,$$

for $\vec{x} = \begin{bmatrix} z_1-\bar{z} \\ z_2-\bar{z} \\ \vdots \\ z_n-\bar{z}\end{bmatrix}$, $\vec{y} = \begin{bmatrix} y_1-\bar{y} \\ y_2-\bar{y} \\ \vdots \\ y_n-\bar{y}\end{bmatrix}$

$-1 \le r \le 1$

5

if $y_i = m x_i$, $m > 0$ then $\vec{y} = m\vec{x}$ (in $\mathbb{R}^n$!), so $r = 1$

if $y_i = m x_i$, $m < 0$, then $r = -1$

if r is close to 1 "strongly positively correlated"

if r is close to -1 "strongly negatively correlated".

Book's example (meat consumption vs. colon cancer in women, by country) is a beautiful example illustrating the difference between correlation ("association") and causation

Correlation (Optional)

Consider the meat consumption (in grams per day per person) and incidence of colon cancer (per 100,000 women per year) in various industrialized countries:

Country	Meat Consumption	Cancer Rate
Japan	26	7.5
Finland	101	9.8
Israel	124	16.4
Great Britain	205	23.3
United States	284	34
Mean	148	18.2

Can we detect a positive or negative *correlation*[4] between meat consumption and cancer rate? Does a country with high meat consumption have high cancer rates, and vice versa? By *high*, we mean "above average," of course. A quick look at the data shows such a *positive correlation:* In Great Britain and the United States, both meat consumption and cancer rate are above average. In the three other countries, they are below average. This positive correlation becomes more apparent when we list the preceding data as *deviations from the mean* (above or below the average):

Country	Meat Consumption (Deviation from Mean) $= \vec{x}$	Cancer Rate (Deviation from Mean) $= \vec{y}$
Japan	−122	−10.7
Finland	−47	−8.4
Israel	−24	−1.8
Great Britain	57	5.1
United States	136	15.8

$$r = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|} \approx \frac{4182.9}{(198.53)(21.4)} \approx .9782$$

Perhaps even more informative is a *scatter plot* of the deviation data. (See Figure 13.)

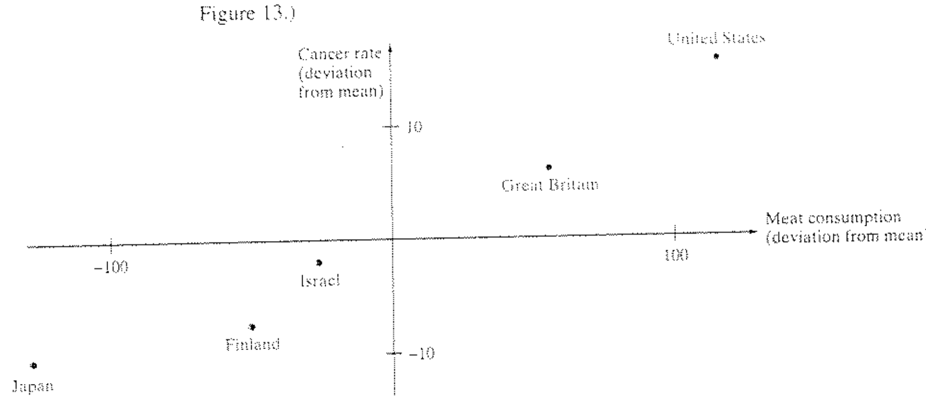


Figure 13

[4] We are using the term *correlation* in a colloquial, *qualitative* sense. Our goal is to *quantify* this term.