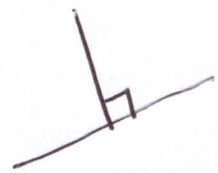
CHAPTER VI: NORM

and

ORTHO-GONALITY

p. 390

This chapter introduces some geometry to vectors, norm or length and orthogonality; "orthogonal" means perpendicular or "making a right angle"  ;

the "square corner" is a traditional drawing.

Orthogonality will be seen to be particularly useful and fundamental, enabling us to minimize bad things such as errors in measurements or models, and to (sometimes) make much easier descriptions such as linear combinations.

Here is a purely algebraic spin on this chapter. Algebra with vectors so far consists only of addition and multiplication by real numbers. Conspicuously missing is multiplication of one vector by another. This chapter will present two ways to do this: the dot product in Sections B through E and the cross product in Section F.

SECTION VIA: NORM

DEFINITION 6.1

Suppose $\vec{x} \equiv (x_1, x_2, \ldots, x_n)$ is a vector in $\mathbb{R}^n$.

The **norm** or **length** or **magnitude** of $\vec{x}$ is

$$\|\vec{x}\| \equiv \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Example 6.2

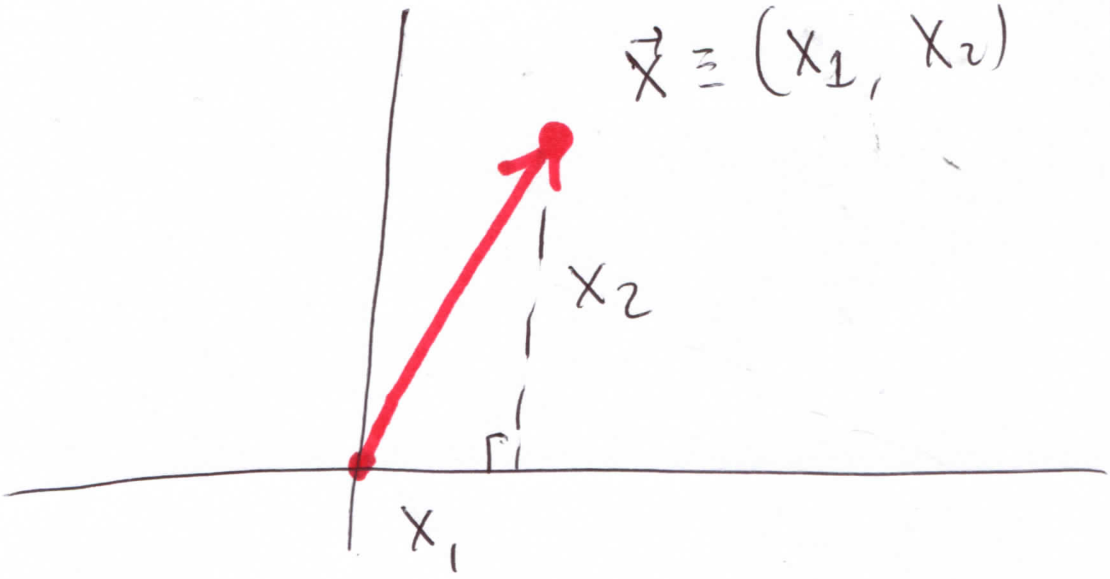
$$\|(1,-2,0,\pi)\| =$$

$$\sqrt{1^2+(-2)^2+0^2+\pi^2} =$$

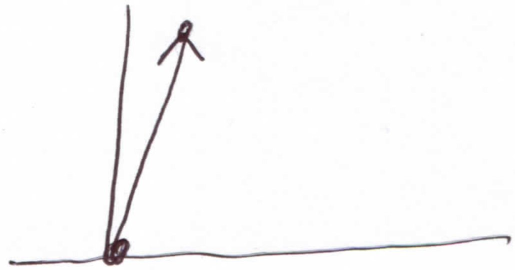
$$\sqrt{5+\pi^2}.$$

REMARKS 6.3

For $n=2$, $\|\vec{x}\|$ is (literally) the length of any directed line segment representing $\vec{x}$, by the Pythagoreon Theorem.



Intuitively, vectors differ from real numbers (sometimes called <u>scalars</u>) in that they have both magnitude <u>and</u> direction. Velocity, such as 50 miles per hour North by Northwest,



is a vector, while speed, such as 50 miles per hour, is a scalar.

If we wish to focus on direction only, it is natural to use vectors whose norm is one.

DEFINITION 6.4

A **unit vector** is a vector whose norm is one.

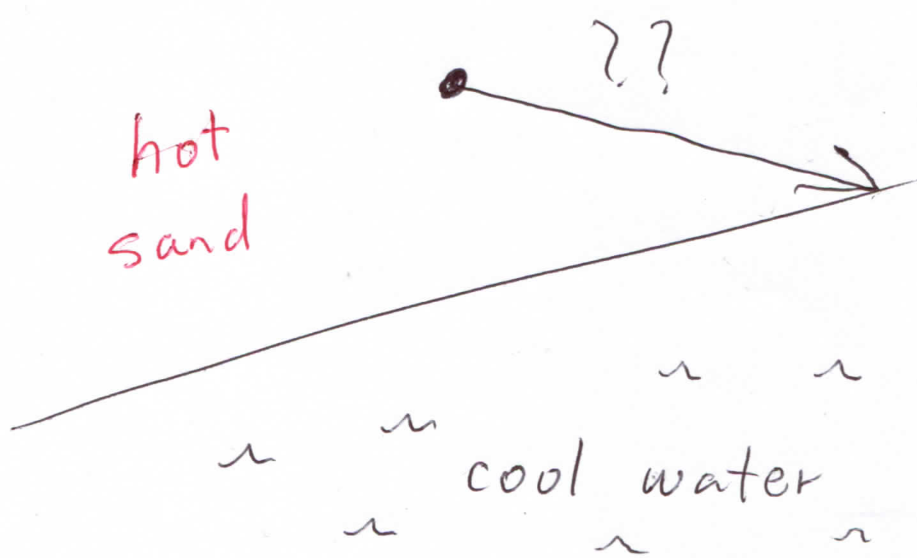
SECTION VIB:

ORTHOGONALITY

and

DOT PRODUCT

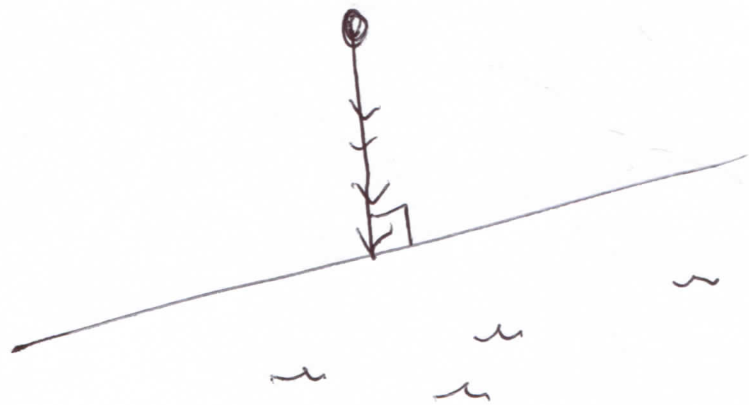
Consider the following scenario.



You're on the beach, with hot sand burning your feet. You rush to the cool water to cool your feet, with your path indicated by the arrow with question marks.

If this were a lecture, we would now ask the class "Why is that path dumb?"

The answer to that insensitive question is that, assuming we want to reach the water as quickly as possible, our path should look like the following:

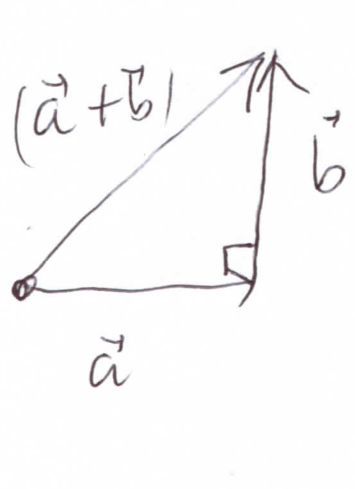


The geometry here is that your path should make a right angle () with the shoreline, to minimize distance.

TERMINOLOGY 6.5

"$\vec{a} \perp \vec{b}$" means the vector $\vec{a}$ is **orthogonal**, or perpendicular, to the vector $\vec{b}$

We'd like to characterize orthogonality algebraically. Motivation comes from the Pythagorean Theorem

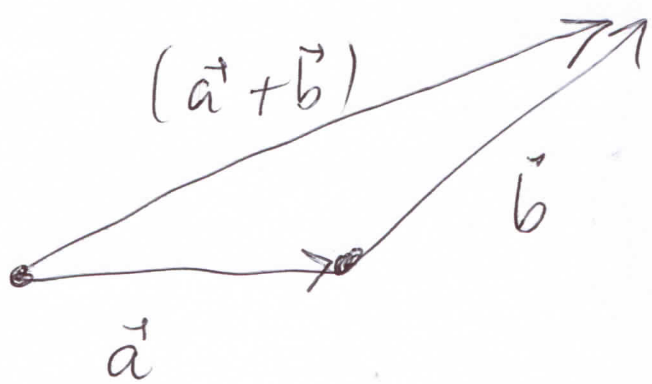


$$\|\vec{a}+\vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2$$

IF $\vec{a} \perp \vec{b}$

The following calculation could be done with vectors in $\mathbb{R}^n$, for any n, but we will do it for $n=2$, for simplicity.

For any vectors $\vec{a} \equiv (a_1, a_2)$ and $\vec{b} \equiv (b_1, b_2)$,



p. 402

$$\|\vec{a}+\vec{b}\|^2 =$$

$$\|(a_1+b_1,\ a_2+b_2)\|^2 =$$

$$(a_1+b_1)^2 + (a_2+b_2)^2 =$$

$$(a_1^2 + 2a_1b_1 + b_1^2) +$$

$$(a_2^2 + 2a_2b_2 + b_2^2) =$$

$$(a_1^2 + a_2^2) + (b_1^2 + b_2^2)$$

$$+ 2(a_1b_1 + a_2b_2) =$$

$$\|\vec{a}\|^2 + \|\vec{b}\|^2 + 2\underbrace{(a_1b_1 + a_2b_2)}_{\text{GIVE THIS a name}}$$

NOTICE that the Pythagorean Theorem holds precisely when that last quantity equals zero.

This motivates the following definition.

DEFINITION 6.6

The **dot** or **inner** or **scalar product** is (In $\mathbb{R}^2$)

$$(a_1, a_2) \bullet (b_1, b_2) \equiv (a_1 b_1 + a_2 b_2).$$

p. 404

Example 6.7

$(1, 2) \bullet (3, -4) = 1 \cdot 3 + 2(-4)$

$= -5.$

DEFINITION 6.8

For any natural number n, if $\vec{a} \equiv (a_1, a_2, \ldots, a_n)$ and $\vec{b} \equiv (b_1, b_2, \ldots, b_n)$, then

$$\vec{a} \bullet \vec{b} \equiv$$

$$(a_1 b_1 + a_2 b_2 + \ldots + a_n b_n)$$

is the **dot** or **inner** or

scalar product

of $\vec{a}$ and $\vec{b}$.

Example 6.9

$$(1, 1, 5, -3) \bullet (4, 7, 0, 2) =$$

$$(1 \cdot 4 + 1 \cdot 7 + 5 \cdot 0 + (-3) \cdot 2) = 5$$

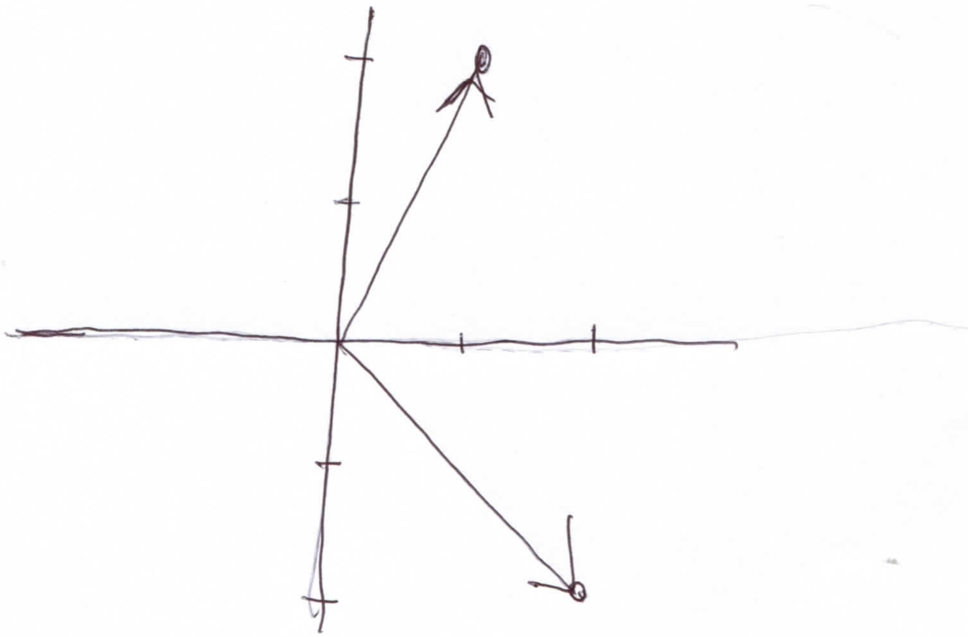
Our calculations before Definition 6.6, generalized to arbitrary n, motivate the following definition.

DEFINITION 6.10

Two vectors $\vec{a}, \vec{b}$ in $\mathbb{R}^n$ are **orthogonal**, $\vec{a} \perp \vec{b}$, if $\vec{a} \cdot \vec{b} = 0$.

Example 6.11 In building or standing, orthogonality is essential for stability or comfort. Staring at vectors may not be sufficient to determine orthogonality, even if we try very hard to make a good picture.

For example, is $(1,2) \perp (2,-2)$? Here's a picture.



Even if we trusted our art skills, we might have trouble deciding if that's a right angle between those vectors.

Algebra relieves us of uncertainty. Take the dot product:

$(1,2) \bullet (2,-2) = 2-4 = -2;$

all we care is that the dot product is <u>not</u> zero, therefore the vectors are <u>not</u> orthogonal.

DOT PRODUCT PROPERTIES 6.12

For n a natural number, $\vec{a}$, $\vec{b}$, and $\vec{c}$ are in $\mathbb{R}^n$ and α is a real number.

(a) $\|\vec{a}\|^2 = \vec{a} \bullet \vec{a}$

(b) $\|\vec{a} + \vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2$
$+ 2(\vec{a} \bullet \vec{b})$

(c) (Pythagorean Theorem)
$\|\vec{a} + \vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2$
if and only if $\vec{a} \perp \vec{b}$.

(d) $\vec{a} \bullet (\vec{b} + \vec{c}) = (\vec{a} \bullet \vec{b}) + (\vec{a} \bullet \vec{c})$

(e) $(\alpha \vec{a}) \bullet \vec{b} = \alpha (\vec{a} \bullet \vec{b})$

(f) $(\vec{a} \bullet \vec{b}) = (\vec{b} \bullet \vec{a})$

(g) $\vec{a} \bullet \vec{a} = 0$ if and only if $\vec{a} = \vec{0}$.

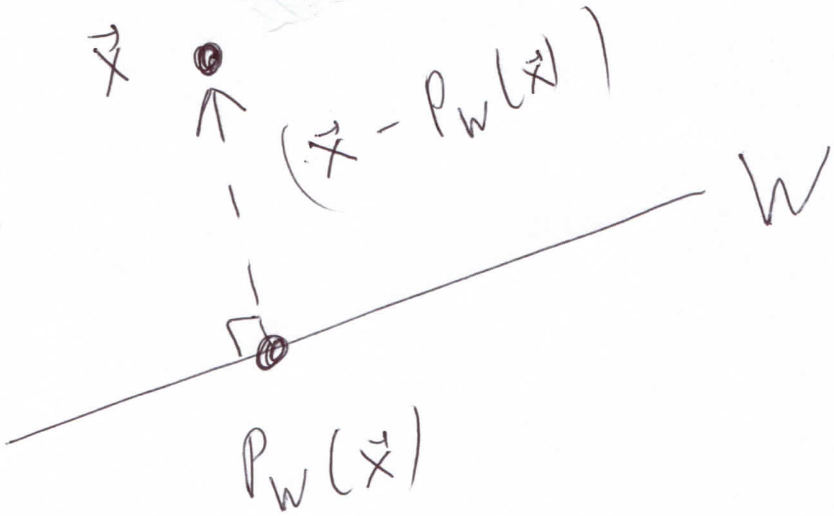
Let's return to the hot-sand-induced motivation at the beginning of this section.

DEFINITION 6.13

If $\vec{x}$ is in $\mathbb{R}^n$ and W is a subspace of $\mathbb{R}^n$, then the **(orthogonal) projection** of $\vec{x}$ onto W, denoted $P_W(\vec{x})$, is a vector in W such that

$(\vec{x} - P_W(\vec{x})) \perp W$; that is

$(\vec{x} - P_W(\vec{x})) \perp \vec{w}$, for all $\vec{w}$ in W



Intuitively, W is reality, $\vec{x}$ is a measurement (mistakes were made), and $P_W(\vec{x})$ is the possible value closest to $\vec{x}$.

p. 412

But let's not take that last assertion on faith.

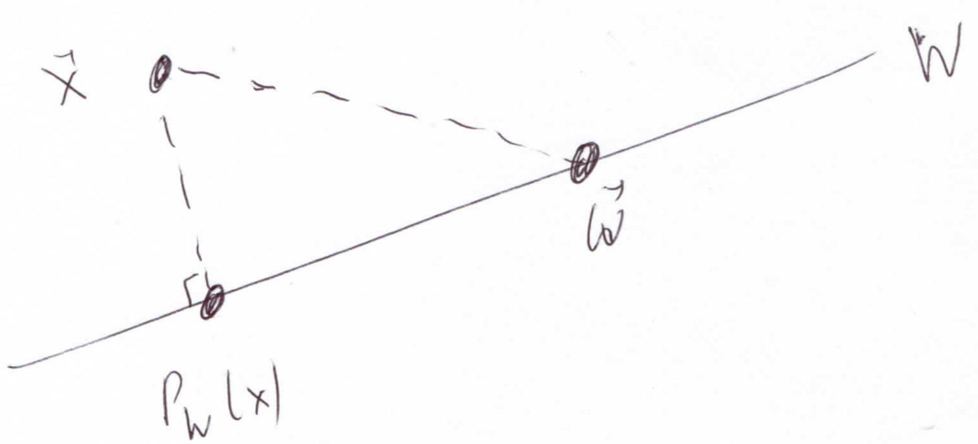
THEOREM 6.14

If $\vec{x}$ is in $\mathbb{R}^n$ and W is a subspace of $\mathbb{R}^n$, then

$$\|\vec{x} - P_W(\vec{x})\| \le \|\vec{x} - \vec{w}\|$$

for all $\vec{w}$ in W.

Proof:



By the Pythagorean theorem,

$$\|\vec{x}-\vec{w}\|^2 = \|(\vec{x}-P_W(\vec{x})) + (P_W(\vec{x})-\vec{w})\|^2$$

$$= \|\vec{x}-P_W(\vec{x})\|^2 + \|P_W(\vec{x})-\vec{w}\|^2$$

$$\geq \|\vec{x}-P_W(\vec{x})\|^2.$$

TERMINOLOGY 6.15

Theorem 6.14 explains why $P_W(\vec{x})$ is called a **best approximation** or **least-squares approximation**

of $\vec{x}$ from W;
"least-squares" refers to the norm of Section VI A.

We will focus for the rest of this section on one-dimensional projections; that is, W in Definition 6.13 equal to a one-dimensional subspace of $\mathbb{R}^n$.

DEFINITION 6.16

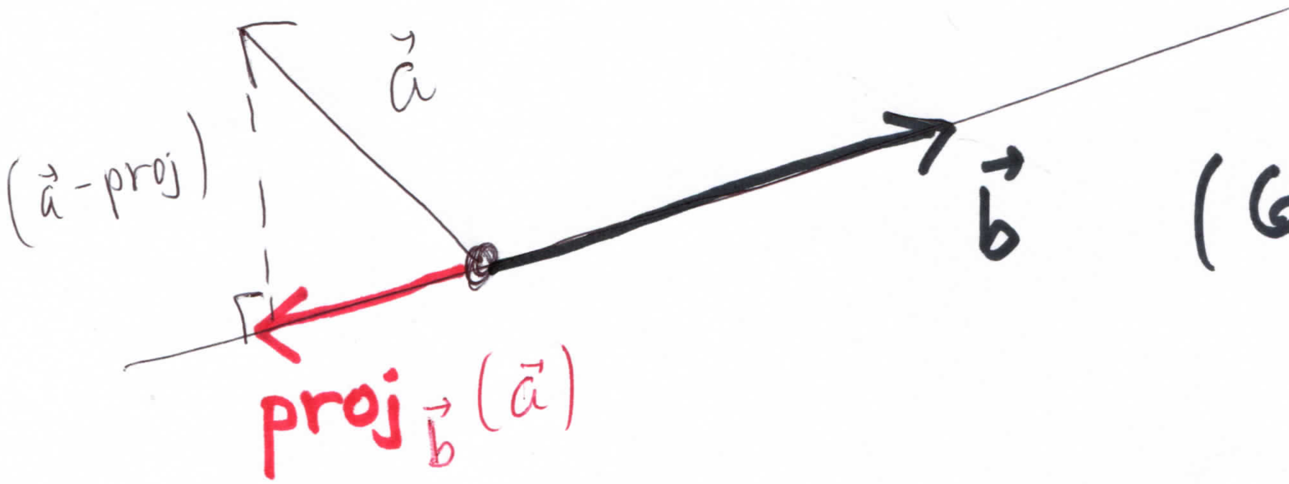
For $\vec{a}, \vec{b}$ vectors in $\mathbb{R}^n$, $\vec{b} \neq \vec{0}$, denote by

$$\text{proj}_{\vec{b}} \equiv P_{\text{span}(\vec{b})};$$

that is, the **(orthogonal) projection of $\vec{a}$ onto $\vec{b}$**, denoted $\text{proj}_{\vec{b}}(\vec{a})$

is a real multiple of $\vec{b}$ such that

$$(\vec{a} - \text{proj}_{\vec{b}}(\vec{a})) \perp \vec{b}.$$



(6.17)

This is a geometric definition; by this we mean that a picture tells the entire story.

As with using the dot product to characterize orthogonality, we would like an algebraic formula for $\text{proj}_{\vec{b}}(\vec{a})$.

First, the definition implies that

$$\text{proj}_{\vec{b}}(\vec{a}) = t\vec{b}$$

for some real t. All we need is t. From (6.17),

$$(\vec{a} - t\vec{b}) \perp \vec{b} \longrightarrow$$

$$0 = (\vec{a} - t\vec{b}) \cdot \vec{b} = (\vec{a} \cdot \vec{b}) - t(\vec{b} \cdot \vec{b}) \longrightarrow$$

$$t = \frac{(\vec{a} \cdot \vec{b})}{\vec{b} \cdot \vec{b}} = \frac{(\vec{a} \cdot \vec{b})}{\|\vec{b}\|^2}.$$

We have shown the following.

Theorem 6.18 If $\vec{a}$ and $\vec{b}$ are in $\mathbb{R}^n$ and $\vec{b} \neq \vec{0}$, then

$$\operatorname{proj}_{\vec{b}}(\vec{a}) = \left[\frac{(\vec{a} \bullet \vec{b})}{\|\vec{b}\|^2}\right]\vec{b}$$

DEFINITION 6.19 The **component** of $\vec{a}$ in the direction of $\vec{b}$ is

$$\operatorname{comp}_{\vec{b}}(\vec{a}) \equiv \frac{\vec{a} \bullet \vec{b}}{\|\vec{b}\|}$$

p. 419

Example 6.20

The projection of $(1,2,3,4)$ onto $(1,0,-1,0)$ is

$$\text{proj}_{(1,0,-1,0)}(1,2,3,4) = \left[\frac{(1,2,3,4)\cdot(1,0,-1,0)}{\|(1,0,-1,0)\|^2}\right](1,0,-1,0)$$

$$= \left(\frac{-2}{2}\right)(1,0,-1,0) = (-1,0,1,0).$$

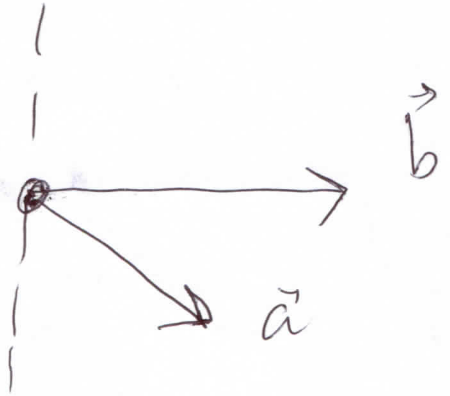
The component of $(1,2,3,4)$ in the direction of $(1,0,-1,0)$ is

$$\text{comp}_{(1,0,-1,0)}(1,2,3,4) = \frac{(1,2,3,4)\cdot(1,0,-1,0)}{\|(1,0,-1,0)\|}$$

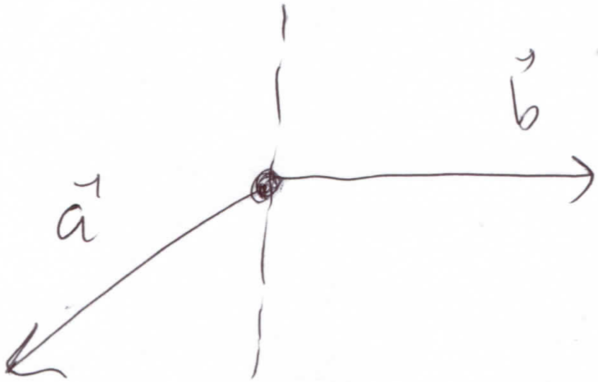
$$= \frac{-2}{\sqrt{2}} = -\sqrt{2}$$

p. 420

Notice that projection is a vector, while component is a number. Component is <u>almost</u> the norm of projection, except that it could be negative.



positive component $\text{comp}_{\vec{b}}(\vec{a})$



negative component $\text{comp}_{\vec{b}}(\vec{a})$

Physically, $\vec{b}$ might be the direction you ride your bike, while $\vec{a}$ is the wind velocity. The only part of $\vec{a}$ that affects your bike speed is $\text{proj}_{\vec{b}}(\vec{a})$; if $\text{comp}_{\vec{b}}(\vec{a})$ is negative, you have some wind in your face; if $\text{comp}_{\vec{b}}(\vec{a})$ is positive, you feel a wind on your bike propelling you forward.

REMARKS 6.21

Results such as Theorem 6.18 and Definition 6.10, that intertwine algebra and geometry, are the best of both worlds: the precision of algebra and the intuition of geometry.

Since we have been talking so much about orthogonality, it is time to mention the other extreme, being parallel.

p. 423

DEFINITIONS 6.22

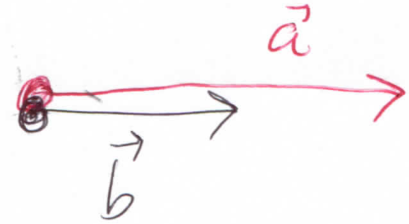
If $\vec{a}$ and $\vec{b}$ are two vectors in $\mathbb{R}^n$, $\vec{a}$ and $\vec{b}$ are **parallel** if one vector is a multiple of the other.

See the vector drawings before Examples 1.22.

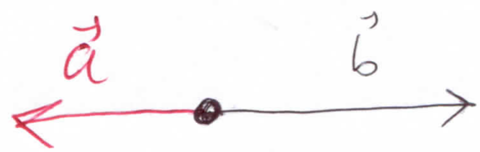
If the multiple above is positive, $\vec{a}$ and $\vec{b}$ **point in the same direction**; if the multiple is negative, $\vec{a}$ and $\vec{b}$ **point in opposite directions.**

p. 424

SAME DIRECTION



OPPOSITE DIRECTION



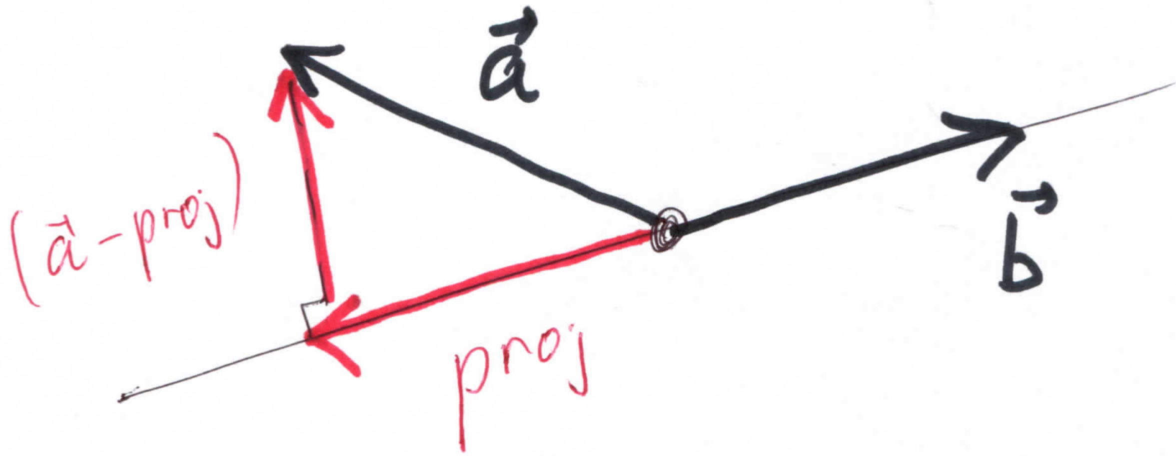
PROPOSITION 6.23

If $\vec{a}$ and $\vec{b}$ are in $\mathbb{R}^n$ and $\vec{b} \neq \vec{0}$, then

$$\vec{a} = \text{proj}_{\vec{b}}(\vec{a}) + \left(\vec{a} - \text{proj}_{\vec{b}}(\vec{a})\right)$$

writes $\vec{a}$ as a sum of two vectors, one parallel to $\vec{b}$, the other orthogonal to $\vec{b}$.

See (6.17).



Example 6.24

Write $(1,0,-1,2,1)$ as a sum of two vectors, one parallel to $(0,1,2,3,4)$, the other perpendicular to $(0,1,2,3,4)$.

SOLUTION:

$$\mathrm{proj}_{(0,1,2,3,4)}(1,0,-1,2,1) =$$

$$\left[\frac{(1,0,-1,2,1)\cdot(0,1,2,3,4)}{\|(0,1,2,3,4)\|^2}\right](0,1,2,3,4)$$

$$= \frac{8}{30}(0,1,2,3,4) = \frac{4}{15}(0,1,2,3,4);$$

$$(1,0,-1,2,1) - \frac{4}{15}(0,1,2,3,4) =$$

$$\frac{1}{15}\left[(15,0,-15,30,15) - (0,4,8,12,16)\right]$$

$$= \frac{1}{15}(15,-4,-23,18,-1); \text{ thus}$$

$$\boxed{(1,0,-1,2,1) = \frac{4}{15}(0,1,2,3,4) + \frac{1}{15}(15,-4,-23,18,-1)}$$

orthogonal (dot product zero)

Since Proposition 6.23 is an orthogonal sum, we may apply the Pythagorean Theorem.

COROLLARY 6.25

For $\vec{a}, \vec{b}$ as in Proposition 6.23,

$$\|\vec{a}\|^2 = \|\vec{a} - \mathrm{proj}_{\vec{b}}(\vec{a})\|^2 + \|\mathrm{proj}_{\vec{b}}(\vec{a})\|^2$$

The following is surprisingly useful, especially for infinite-dimensional dot products.

p. 428

CAUCHY - SCHWARZ INEQUALITY 6.26

(In finite dimension)

For $\vec{a}, \vec{b}$ in $\mathbb{R}^n$,

$$|\vec{a} \cdot \vec{b}| \leq \|\vec{a}\| \|\vec{b}\|.$$

Proof: If $\vec{b} = \vec{0}$, both sides of the inequality are zero

If $\vec{b} \neq \vec{0}$, then by Corollary 6.25, $\|\vec{a}\| \geq \|\text{proj}_{\vec{b}}(\vec{a})\| =$

$$\left\| \left(\frac{\vec{a} \cdot \vec{b}}{\|\vec{b}\|^2} \right) \vec{b} \right\| = \frac{|\vec{a} \cdot \vec{b}|}{\|\vec{b}\|}.$$

REMARKS and Example 6.27.

If armed with trigonometry, the Cauchy-Schwarz inequality allows us to calculate the measures of angles between arbitrary nontrivial vectors with dot products; see Appendix Four.

For integral powers of integers k, there are many formulas for sums of powers of k; e.g.,

$$(1+2+3+\dots+n) = \frac{n(n+1)}{2},$$

$$(1^2+2^2+3^2+\dots+n^2) = \frac{n(n+1)(2n+1)}{6}, \text{ etc.}$$

For nonintegral powers, such formulas might not exist, but Cauchy-Schwarz allows us to estimate; e.g.

$$(\sqrt{1}+\sqrt{2}+\sqrt{3}+\dots+\sqrt{n}) =$$

$$(1,1,\dots,1)\cdot(\sqrt{1},\sqrt{2},\sqrt{3},\dots,\sqrt{n})$$

$$\leq \|(1,1,\dots,1)\|\,\|(\sqrt{1},\sqrt{2},\dots,\sqrt{n})\|$$

$$= \sqrt{n}\sqrt{1+2+3+\dots+n}$$

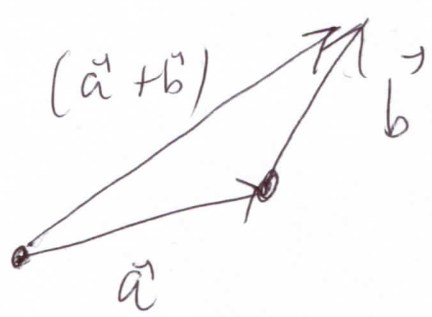
$$= \sqrt{n}\sqrt{\frac{n(n+1)}{2}} = n\frac{\sqrt{n+1}}{\sqrt{2}}.$$

Another consequence of Cauchy Schwarz is the following, which is very believable when one draws a picture of vector addition.

TRIANGLE INEQUALITY 6.28

$$\|\vec{a}+\vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\|,$$

for any $\vec{a}, \vec{b}$ in $\mathbb{R}^n$.



p. 432

Proof: By 6.12 and Cauchy Schwarz, in that order,

$$\|\vec{a}+\vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\| + 2(\vec{a}\cdot\vec{b})$$

$$\leq \|\vec{a}\|^2 + \|\vec{b}\|^2 + 2\|\vec{a}\|\,\|\vec{b}\|$$

$$= (\|\vec{a}\| + \|\vec{b}\|)^2$$