Shifrin 1.2. The Dot Product

Recall: $\mathbf{x} \in \mathbb{R}^n$ is given by $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

$$\mathbf{x} \pm \mathbf{y} = \begin{bmatrix} x_1 \pm y_1 \\ \vdots \\ x_n \pm y_n \end{bmatrix} \qquad c\mathbf{x} = \begin{bmatrix} cx_1 \\ \\ cx_n \end{bmatrix}$$

Definition. The <u>dot product</u> or <u>inner product</u> of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is the <u>scalar</u>

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \dots + x_n y_n$$

This is a kind of multiplication for vectors (but be careful - the output is <u>not</u> a vector), and it has similar properties:

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ ← commutative
2. $\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2 \geq 0$, $\|\mathbf{x}\| = 0 \iff \mathbf{x} = 0$
3. $c(\mathbf{x} \cdot \mathbf{y}) = (c\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (c\mathbf{y})$
4. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$

   ↑ distributes over vector addition

We can use these properties to do vector algebra that looks much like the ordinary (scalar) algebra you already know.

Lemma. $\|x+y\|^2 = \|x\|^2 + 2x\cdot y + \|y\|^2$

Proof.
$$\begin{aligned}\|x+y\|^2 &= (x+y)\cdot(x+y)\\ &= x\cdot x + x\cdot y + y\cdot x + y\cdot y\\ &= x\cdot x + 2x\cdot y + y\cdot y\\ &= \|x\|^2 + 2x\cdot y + \|y\|^2. \quad \square\end{aligned}$$

Definition. We say $x, y \in \mathbb{R}^n$ are *orthogonal* or *perpendicular* if $x\cdot y = 0$.

Definition. Given $x, y \in \mathbb{R}^n$, we let

$$x^{\parallel} = \frac{x \cdot y}{\|y\|^2} y, \quad x^{\perp} = x - \frac{x \cdot y}{\|y\|^2} y$$

Proposition. We have

1. $x^{\parallel}$ is parallel to y
2. $x^{\perp}$ is orthogonal to y
3. $x = x^{\parallel} + x^{\perp}$

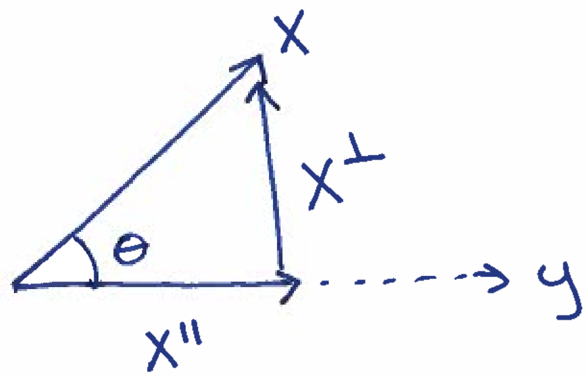
Proof. The only one that's not easy is 2.

$$x^{\perp} \cdot y = \left(x - \frac{x \cdot y}{\|y\|^2} \right) \cdot y$$

$$= x \cdot y - \frac{x \cdot y}{\cancel{\|y\|^2}} \cancel{y \cdot y}$$

$$= 0. \qquad \square$$

Now suppose we have (in $\mathbb{R}^2$)



we Know

adjacent side

$$\cos\theta = \frac{\|x^\parallel\|}{\|x\|} = \frac{\left\| \frac{x\cdot y}{\|y\|^2}\, y \right\|}{\|x\|}$$

hypoteneuse

$$= \frac{x\cdot y}{\|y\|^{\cancel{2}}} \, \frac{\cancel{\|y\|}}{\|x\|}$$

so

$$= \frac{x\cdot y}{\|x\|\,\|y\|}$$

Lemma. In $\mathbb{R}^2$, the angle θ between x and y is given by

$$\theta = \arccos\left(\frac{x\cdot y}{\|x\|\,\|y\|} \right).$$

<u>Definition</u>. The angle between $x, y \in \mathbb{R}^n$ is the unique θ in $[0, \pi]$ so that

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

Now we ~~can~~ show something <u>important</u>.

Cauchy-Schwarz Inequality.

If $x, y \in \mathbb{R}^n$, then $|x \cdot y| \leq \|x\| \, \|y\|$.
Further, $|x \cdot y| = \|x\| \, \|y\|$ if and only if one vector is a scalar multiple of the other.

Proof. If $y = 0$ then $|x \cdot y| = 0 = \|x\| \, \|y\|$ and $y = 0x$, so we're done.

scalar 0!

If $y \neq 0$, we consider

$$g(t) = \|x + ty\|^2 = \|x\|^2 + 2t \, x \cdot y + t^2 \|y\|^2$$

Now $g(t)$ is a square (and a norm) so $g(t) \geq 0$ for all t. At the min,

$$g(t_0) = \|x\|^2 - 2\frac{(x\cdot y)^2}{\|y\|^2} + \frac{(x\cdot y)^2}{\|y\|^{\cancel{4}\,2}} \cancel{\|y\|^2}$$

$$= \|x\|^2 - \frac{(x\cdot y)^2}{\|y\|^2}$$

so

$$\|x\|^2 - \frac{(x\cdot y)^2}{\|y\|^2} \geq 0$$

~~and~~

$$\|x\|^2 \geq \frac{(x\cdot y)^2}{\|y\|^2}$$

$$\|x\|^2 \|y\|^2 \geq (x\cdot y)^2$$

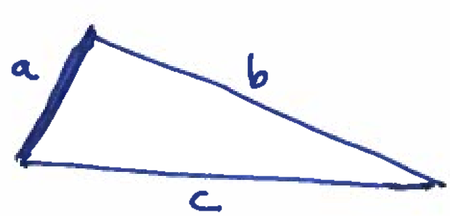
and (taking square roots)

$$\|x\|\,\|y\| \geq |x\cdot y|$$

Now $\|x\|\,\|y\| = |x\cdot y| \overset{\text{if and only if}}{\Longleftrightarrow} g(t_0)=0$, or $\|x - t_0 y\|^2 = 0$. In this case $\mathbf{x} = t_0 \mathbf{y}$ and $\mathbf{x}$ is a scalar multiple of $\mathbf{y}$. $\square$

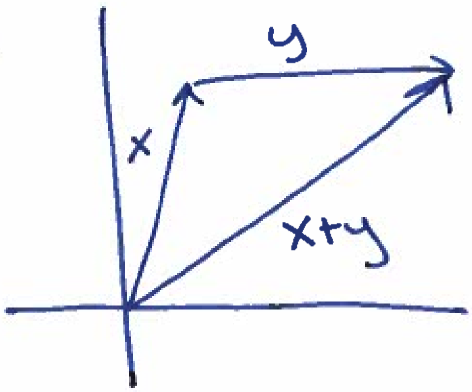
We can now use this to prove.

<u>Triangle Inequality</u>. The sum of the lengths of two sides of a triangle is $\geq$ the length of the third side.



$a+b \geq c$
$b+c \geq a$
$c+a \geq b$

Proof.



If we write ΔABC in vector form, we must show $\|x+y\| \leq \|x\| + \|y\|$.

But

$$\|x+y\|^2 = \|x\|^2 + 2x \cdot y + \|y\|^2$$

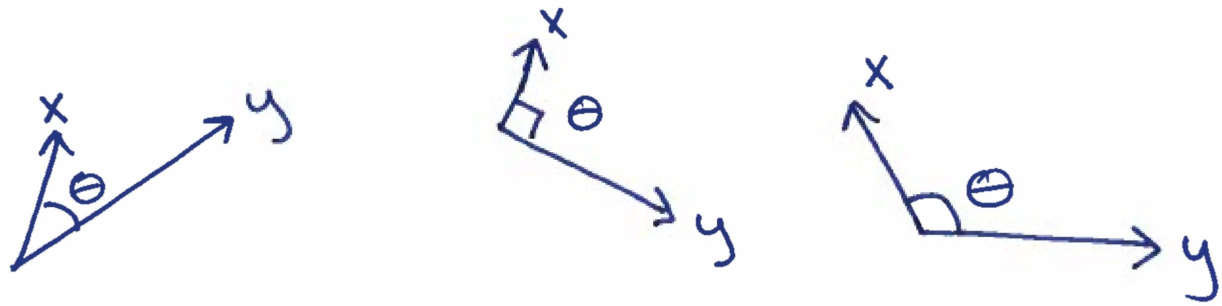
$$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$$

$$= (\|x\| + \|y\|)^2$$

So $\|x+y\| \leq \|x\| + \|y\|$. $\square$

Note: We used the fact that all these norms are ≥ 0. Where?

Examples.



$\theta < \pi/2$ $\qquad$ $\theta = \pi/2$ $\qquad$ $\theta > \pi/2$

$x \cdot y > 0$ $\qquad$ $x \cdot y = 0$ $\qquad$ $x \cdot y < 0$

Suppose we define two vectors by data about n people.

$$\vec{x} = \begin{bmatrix} \vdots \\ x_i \\ \vdots \end{bmatrix} \qquad x_i = \begin{cases} 1, & \text{if student } i \text{ is wearing hat} \\ -1, & \text{if not} \end{cases}$$

$$\vec{y} = \begin{bmatrix} \vdots \\ y_i \\ \vdots \end{bmatrix} \qquad y_i = \begin{cases} 1, & \text{if student } i \text{ is wearing sunglasses} \\ -1, & \text{if not} \end{cases}$$

Suppose that $\mathbb{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$.

Definition. The <u>mean</u> or <u>average</u> of $X \in \mathbb{R}^n$ is given by $\mu_X = \frac{1}{n} X \cdot \mathbb{1} = \frac{1}{n} \sum X_i$.

We can "zero mean" X by taking the component of X orthogonal to $\mathbb{1}$.

$$X^{\perp} = X - \frac{X \cdot \mathbb{1}}{\|\mathbb{1}\|^2} \mathbb{1}$$

$$= X - \frac{X \cdot \mathbb{1}}{n^2} \mathbb{1}$$

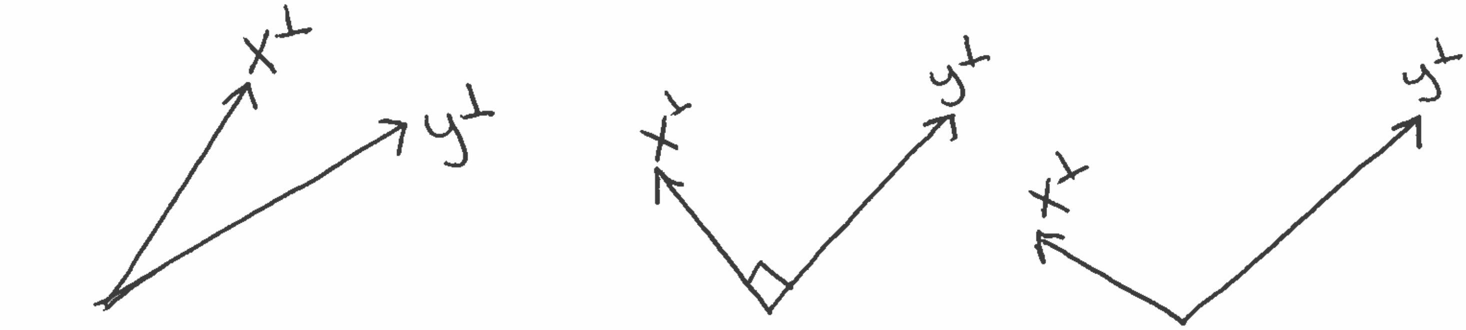
$$= X - \left(\frac{1}{n} X \cdot \mathbb{1}\right) \frac{1}{n} \mathbb{1}$$

$$= X - \mu_X \cdot \frac{1}{n} \mathbb{1}$$

Definition. The Pearson correlation coefficient (or r) between x and y is

$$r = \frac{x^\perp \cdot y^\perp}{\|x^\perp\| \, \|y^\perp\|} = \cos\theta$$

where θ is the angle between $x^\perp, y^\perp$.



positively correlated or "correlated" | uncorrelated | negatively correlated or "anticorrelated"

We will keep returning to this picture

statistics $\longleftrightarrow$ high dimensional geometry