3/5/2019: Norms.

The absolute value function is a special example of a norm on the vector space of real numbers, $V = \mathbb{R}$.

$$|\cdot| : \mathbb{R} \to \mathbb{R}^+ \cup \{0\}$$ ~~satisfying~~

defined as: $$|v| = \begin{cases} v & \text{if } v \geq 0. \\ -v & \text{if } v < 0. \end{cases}$$

Properties:

- $\rightarrow$ $|v| \geq 0 \quad \forall v \in \mathbb{R}$, $|v| = 0 \iff v = 0$.
- $\rightarrow$ $|\lambda v| = |\lambda|\,|v| \quad \forall \lambda \in \mathbb{R}$ and $\forall v \in \mathbb{R}$
- $\rightarrow$ $|u+v| \leq |u| + |v| \quad \forall u, v \in \mathbb{R}$.

We used this norm to measure convergence of sequences. For example, we know the sequence of iterates from Newton's method satisfies:

$$|x_{k+1} - \xi| \leq C\,|x_k - \xi|^2.$$

i.e. a norm gives us a measure of distance in a vector space V.

Definition 2.6 S'pose V is a vector space. (What does this mean?). V is closed under addition and scalar multiplication. Here, are scalars are just real numbers. We say $\|\cdot\|$ is a *Norm* on V if it satisfies:

1. $\|\underline{v}\| = 0 \iff \underline{v} = 0$ in V
2. $\|\lambda \underline{v}\| = |\lambda| \, \|\underline{v}\| \quad \forall \lambda \in \mathbb{R}$ and $\forall \underline{v} \in V$
3. $\|\underline{u} + \underline{v}\| \leq \|\underline{u}\| + \|\underline{v}\| \quad \forall \underline{u}, \underline{v} \in V$.

The ~~thing~~ pair $(V, \|\cdot\|)$ is called a *normed linear space*, or a *normed vector space*.

Some examples of norms on the vector space $V = \mathbb{R}^n$, given as definitions since these norms are important.

Definition 2.7: "1-Norm"

$$\|\underline{v}\|_1 = \sum_{i=1}^{n} |v_i|.$$

Definition 2.8 "2-Norm"

$$\|\underline{v}\|_2 = \left(\sum_{i=1}^{n} v_i^2\right)^{1/2}$$

Note that you can write the 2-norm as

$$(\underline{v} \cdot \underline{v})^{1/2} = \|\underline{v}\|_2 .$$

i.e. this norm is "induced" from an inner product on $\mathbb{R}^n$.

Definition 2.9 The "∞-norm" is

$$\|\underline{v}\|_\infty = \max_{i=1,\dots,n} |v_i| .$$

Rmk: when $n=1$, i.e. $V = \mathbb{R}$, then

$$\|\underline{v}\|_1 = \|\underline{v}\|_2 = \|\underline{v}\|_\infty = |\underline{v}| = |v| .$$

Example: let $V = \mathbb{R}^3$, $\underline{v} = (1, -2, 3.14)$.

$$\|\underline{v}\|_1 = 1 + 2 + 3.14 = 6.14 .$$

$$\|\underline{v}\|_2 = 1 + 4 + 3.14^2 = 14.8596$$

$$\|\underline{v}\|_\infty = \max\{|1|, |-2|, |3.14|\} = 3.14 .$$

Very important:

Cauchy-Schwartz Inequality: $\left|\sum_{i=1}^n u_i v_i\right| \le \|\underline{u}\|_2 \|\underline{v}\|_2 \quad \forall \underline{u}, \underline{v} \in \mathbb{R}^n .$

Cauchy-Schwarz is a tool useful for many many things. In particular, we will use it to show the triangle inequality.

Proof of Cauchy-Schwarz

Take $\lambda \in \mathbb{R}$, $\underline{u}, \underline{v} \in \mathbb{R}^n$ arbitrary.

$$0 \leq \| \lambda \underline{u} + \underline{v} \|_2^2 = \sum_{i=1}^{n} (\lambda u_i + v_i)^2$$

$$= \lambda^2 \underbrace{\sum_{i=1}^{n} |u_i|^2}_{=A} + 2\underbrace{\lambda \sum_{i=1}^{n} u_i v_i}_{=B} + \underbrace{\sum_{i=1}^{n} |v_i|^2}_{=C}.$$

We have a nonnegative quadratic polynomial in λ:

$$f(\lambda) = A\lambda^2 + B\lambda + C \geq 0.$$

So, either $f(\lambda)$ has only one root:

$$B^2 - 4AC = 0$$

or it has no real roots, i.e.

$$B^2 - 4AC < 0.$$

→

So we have in either case.

$$B^2 - 4AC \le 0$$

$$\Leftrightarrow \left(2 \sum_{i=1}^{n} u_i v_i\right)^2 - 4\left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right) \le 0$$

$$\Leftrightarrow \left(\sum_{i=1}^{n} u_i v_i\right)^2 \le \left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right).$$

Taking the square root of both sides gives:

$$\left(\sum_{i=1}^{n} u_i v_i\right) \le \|\underline{u}\|_2 \ \|\underline{v}\|_2 .$$

<u>Remark:</u> Equality holds in Cauchy-Schwarz if and only if ($\Leftrightarrow$) $\underline{u} = \lambda \underline{v}$ for some $\lambda \in \mathbb{R}$, $\lambda \geq 0$.

($\Leftarrow$) Easier direction: let $\underline{u} = \lambda \underline{v}$.

Then $\|\underline{u}\| \, \|\underline{v}\| = \|\lambda \underline{v}\| \, \|\underline{v}\| = |\lambda| \, \|\underline{v}\|^2$

$$= |\lambda| \, \underline{v} \cdot \underline{v} = \lambda \underline{v} \cdot \underline{v} = \underline{u} \cdot \underline{v}.$$

($\Rightarrow$) Assume equality i.e. $\underline{u} \cdot \underline{v} = \|\underline{u}\| \|\underline{v}\|$.

Note that $\|\underline{u}\| \|\underline{v}\| \cos\theta = \underline{u} \cdot \underline{v}$,

with θ = angle between vectors $\underline{u}$ and $\underline{v}$.

Assume both $\underline{u}$ and $\underline{v}$ are not the zero vector, for if that were true then the result trivially holds. then

$$\underline{u} \cdot \underline{v} = \|\underline{u}\| \|\underline{v}\| \Leftrightarrow \|\underline{u}\| \|\underline{v}\| \cos\theta$$

$$= \|\underline{u}\| \|\underline{v}\|$$

$$\Leftrightarrow \cos\theta = 1,$$

So $\theta = 0$, i.e. vectors $\underline{u}$ and $\underline{v}$ are parallel. So $\underline{u} = \lambda \underline{v}$, for some $\lambda \in \mathbb{R}$, $\lambda \geq 0$.

Now, we can use Cauchy-Schwarz to prove the triangle inequality for the 2-Norm:

$$\|\underline{u}+\underline{v}\|_2^2 = \|\underline{u}\|_2^2 + 2\sum_{i=1}^{n} u_i v_i \quad \|\underline{v}\|_2^2$$

$$\begin{pmatrix}\text{Cauchy}\\ \text{Schwarz}\end{pmatrix} \leq \|\underline{u}\|_2^2 + 2\|\underline{u}\|_2\|\underline{v}\|_2 + \|\underline{v}\|_2^2$$

$$= \left(\|\underline{u}\|_2 + \|\underline{v}\|_2\right)^2$$

implying $\|\underline{u}+\underline{v}\|_2 \leq \|\underline{u}\|_2 + \|\underline{v}\|_2$.

<u>Definition</u>: "p-Norm." Take $p \geq 1$.

it is defined as

$$\|\underline{v}\|_p = \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p} \quad \forall\, \underline{v} \in \mathbb{R}^n.$$

How do we show the triangle inequality in general for the p-Norm?

In the process of showing that $\|\cdot\|_p$ satisfies the triangle inequality, we need some tools which are important in their own right.

<u>Thm 2.4</u> (Young's Inequality)

Let $p, q > 1$, and $\frac{1}{p} + \frac{1}{q} = 1$.

Then, for any $a, b \geq 0$, we have

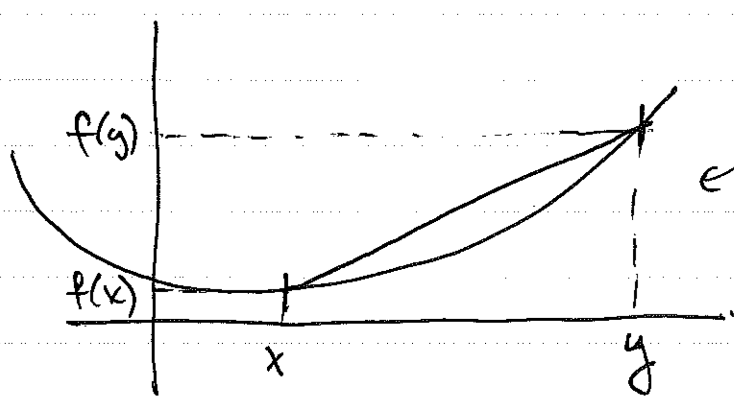
$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

<u>Proof:</u> Assume $a > 0$ and $b > 0$. If either $a = 0$ or $b = 0$, the result is trivial.

Recall the definition of convexity of a function: $x \mapsto f(x)$ is convex if

$$f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y).$$

$\forall\ \theta \in [0,1]$, $\forall x, y \in \mathbb{R}$.



graph of f lies below the "chord" connecting $f(x)$ and $f(y)$.

Notice that $x \mapsto e^x$ is convex. Doing some manipulations, using the property of logs, we have:

$$ab = e^{\log(ab)}$$

$$= e^{\log a + \log b}$$

$$= e^{(\frac{1}{p})\log a^p + (\frac{1}{q})\log b^q}$$

Now, apply convexity of e^x with $\theta = \frac{1}{p}$, $(1-\theta) = 1 - \frac{1}{p} = \frac{1}{q}$, $x = \log a^p$, $y = \log b^q$, to conclude

$$e^{(\frac{1}{p})\log a^p + (\frac{1}{q})\log b^q}$$

$$\leq \frac{1}{p} e^{\log a^p} + \frac{1}{q} e^{\log b^q}$$

$$= \frac{a^p}{p} + \frac{b^q}{q},$$

So we can conclude:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Thm: 2.5 (Hölder's inequality).

let $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. Then

$\forall\ \underline{u}, \underline{v} \in \mathbb{R}^n$ we have:

$$\left| \sum_{i=1}^{n} u_i v_i \right| \leq \|\underline{u}\|_p \|\underline{v}\|_q .$$

Proof: Assume $\underline{u} \neq 0$ and $\underline{v} \neq 0$, otherwise the inequality trivially holds. Construct normalized versions of $\underline{u}$ and $\underline{v}$:

$$\tilde{\underline{u}} = \frac{\underline{u}}{\|\underline{u}\|_p}, \quad \tilde{\underline{v}} = \frac{\underline{v}}{\|\underline{v}\|_q} .$$

Now, we look at:

$$\left| \sum_{i=1}^{n} \tilde{u}_i \tilde{v}_i \right| \leq \sum_{i=1}^{n} |\tilde{u}_i \tilde{v}_i| \quad \left(\text{triangle inequality}\right).$$

$$\left(\text{Young's inequality}\right) \leq \frac{1}{p} \sum_{i=1}^{n} |\tilde{u}_i|^p + \frac{1}{q} \sum_{i=1}^{n} |\tilde{v}_i|^q$$

$$= \frac{1}{p} + \frac{1}{q} = 1.$$

So, we've shown that

$$\left| \sum_{i=1}^{n} \tilde{u}_i \tilde{v}_i \right| \leq 1.$$

Plugging in $\tilde{u}_i = \dfrac{u_i}{\|u\|_p}$, $\tilde{v}_i = \dfrac{v_i}{\|v\|_q}$,

we conclude $\left| \sum_{i=1}^{n} u_i v_i \right| \leq \|u\|_p \|v\|_q$.

Thm 2.6: (Minkowski's Inequality).

Let $1 \leq p \leq \infty$ and $\underline{u}, \underline{v} \in \mathbb{R}^n$. Then

$$\|\underline{u} + \underline{v}\|_p \leq \|\underline{u}\|_p + \|v\|_p.$$

Proof: $p=1$, $p=\infty$ are easy. (☺)

Assume that $1 < p < \infty$. Also assume that $\underline{u} \neq 0$ and $\underline{v} \neq 0$. Then:

$$\|\underline{u} + \underline{v}\|_p^p = \sum_{i=1}^{n} |u_i + v_i|^p$$

$$= \sum_{i=1}^{n} |u_i + v_i|^{p-1} |u_i + v_i|$$

$$\leq \sum_{i=1}^{n} |u_i+v_i|^{p-1} (|u_i| + |v_i|) \quad \left(\begin{array}{c}\text{Triangle}\\ \text{inequality}\end{array}\right).$$

Note that By Hö"lder's inequality, we can write:

$$\sum_{i=1}^{n} |u_i+v_i|^{p-1} |u_i| \leq \|(\underline{u}+\underline{v})^{p-1}\|_{\frac{p}{p-1}} \quad \|\underline{u}\|_p$$

$$= \left(\sum_{i=1}^{n} (u_i+v_i)^p\right)^{\frac{p-1}{p}} \left(\sum_{i=1}^{n} |u_i|^p\right)^{1/p}.$$

$$= \|\underline{u}+\underline{v}\|_p^{p-1} \|\underline{u}\|_p .$$

Thus, we can continue:

$$\|\underline{u}+\underline{v}\|_p^p \leq \|\underline{u}+\underline{v}\|_p^{p-1} \left(\|\underline{u}\|_p + \|v\|_p\right).$$

Dividing through by $\|\underline{u}+\underline{v}\|_p^{p-1}$, we obtain the inequality.

Picture for Norms:

Consider the sets of vectors in $\mathbb{R}^2$ defined as follows:

$$S'_1 = \{ \underline{v} \in \mathbb{R}^2 : \|\underline{v}\|_1 = 1 \}.$$

$$S'_2 = \{ \underline{v} \in \mathbb{R}^2 : \|\underline{v}\|_2 = 1 \}.$$

$$S'_\infty = \{ \underline{v} \in \mathbb{R}^2 : \|\underline{v}\|_\infty = 1 \}.$$