2 Norms

Problem with Scientific Computing: The answer always is always wrong.

Natural question: How wrong?

Need a measure of distance.

Actually: A measure of the length of a vector is fine, too. Why?

Can simply measure the length of a difference.

Notation: Length a vector written as $\|\vec{x}\|$

Actually called the "norm" of the vector

A little like an absolute value for vectors

To measure distance, we use norm (length) of the difference:

$$\text{distance}(\vec{x},\vec{y}) = \|\vec{x}-\vec{y}\|$$

Most commonly, we will measure (or predict) the norm of an error:

correct answer: $\vec{u}$

computed answer: $\vec{u}_{approx}$

difference = error: $\vec{u}_{approx} - \vec{u}$

error norm: $\|\vec{u}_{approx} - \vec{u}\|$

small: good
large: bad

So how do we compute a 'norm'?

Well, there's more than one way to do it... as always. :)

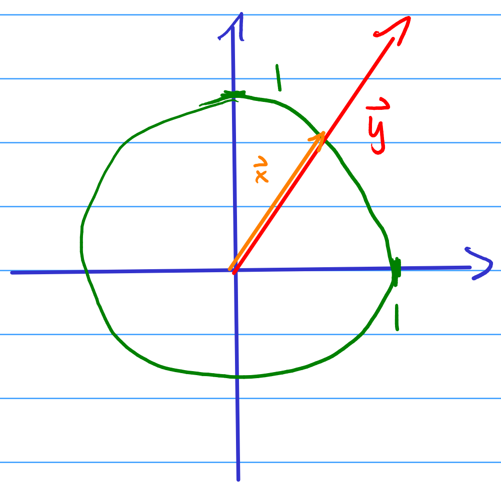
Say we assume that the norm of the vector $2\vec{x}$ has twice the norm of $\vec{x}$ (and similarly for other factors).

Then all we need to know is all the vectors with norm 1.

Those are important, we'll call them unit vectors (relative to our norm).

We can then find the length of any vector just by multiplying.

Get to the point.



What is the norm of this vector?

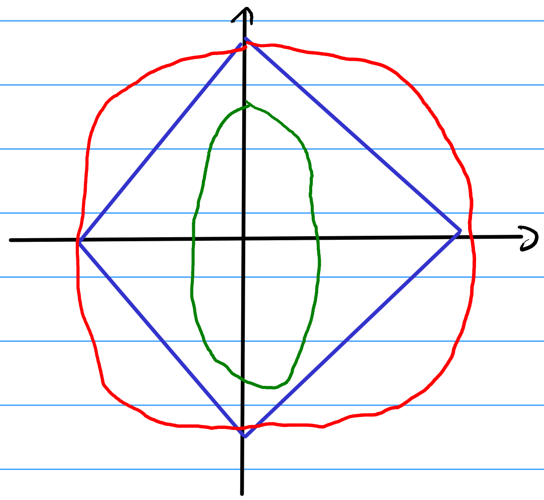
$\|\vec{x}\| = 1$

$\vec{y} = 1.8\,\vec{x}$

$\leadsto \|\vec{y}\| = 1.8$

Suppose these are all our unit vectors.

Is the circle the only possible unit ball (in two dimensions)?



No!

So, what does a norm have to satisfy to call itself a norm?

- $\|x\| \geq 0$
- $\|x+y\| \leq \|x\| + \|y\|$ (the "triangle inequality")
- $\|\alpha x\| = |\alpha| \, \|x\|$
- $\|x\| = 0 \Leftrightarrow x = 0$

To think about: Why is each of these important?

Examples! I need examples!

One common example are the so-called p-norms on vectors in $\mathbb{R}^n$.

Assume $p \geq 1$.

Then for $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$$\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}.$$

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

$$\|x\|_1 = |x_1| + \cdots + |x_n|$$

$$\|x\|_\infty \text{"="} \sqrt[\infty]{|x_1|^\infty + |x_2|^\infty + \cdots + |x_n|^\infty}.$$

$$= \max(|x_1|, |x_2|, \ldots, |x_n|)$$

Demo: Compute p-norms of a vector for increasing p.

Show evolution towards ∞-norm.

In principle, we would need to verify that these norms satisfy our properties. We will simply take them for granted however.

Now we've put together this machinery of 'norms'. How does that help us?

Typically our inputs and outputs are vectors of some sort.

We model all our data (both inputs and outputs of a computation) as:

Data = True Value + Error

$$\vec{x} = \vec{x}_0 + \vec{\Delta x}$$

$$\rightsquigarrow \vec{\Delta x} = \vec{x} - \vec{x}_0$$

$$\|\vec{\Delta x}\| = \|\vec{x} - \vec{x}_0\|$$

So we could say, "The norm of your error is about 25.3."

This is called absolute error.

But it's not enough to say whether a result is *good*.

Examples:

If the norm of $\vec{x}_0$ is 15,000, then that would be 0.16% error -> probably good.

If the norm of $\vec{x}_0$ is 30, then the error is about as big as the result -> less good.

Also define relative error:

$$\frac{\|\vec{\Delta x}\|}{\|\vec{x}_0\|} = \frac{\|\vec{x} - \vec{x}_0\|}{\|\vec{x}_0\|}$$

Example: Relative error = 0.01 means that we're only off by one percent.

We will study ways to solve "problems".

Problems have an input (data) and an output (answer) . Both inaccurate (have error).

We want to say something about error at output vs error at input.

What could we say?

We could ask how much the error gets amplified by the method, i.e.

$$\text{(Relative) Condition number} = \max_{\text{over all inputs}} \frac{\text{(Relative) error in output}}{\text{(Relative) error in input}}$$

⚠ The condition number is a property of the problem

-> Cannot depend on the input!

Is a condition number typically < 1?

No, because then our output *always* has smaller relative error than the input. That doesn't happen very often.

Can we take norms of matrices, too?

The short answer is "Yes, of course--you can just read the numbers in the matrix as a vector and apply a vector norm."

But that falls short.

Why does that fall short?

Imagine a matrix and a vector along with a norm $\|\cdot\|$.

The "dot" notation means:
'argument of the norm goes here'.

Also imagine we have a norm $|||\cdot|||$ for matrices.

It seems reasonable to ask that $\|Ax\| \le |||A|||\,\|x\|$.

So, given a vector norm $\|\cdot\|$, we can use this relationship to define the corresponding matrix norm:

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

Then $\|Ax\| = \left\|A \frac{x}{\|x\|} \cdot \|x\|\right\| = \left\|A \underbrace{\frac{x}{\|x\|}}_{\|\cdot\|=1}\right\| \cdot \|x\| \le \|A\|\,\|x\|.$

Demo: Finding matrix norm values

In the next chapter, we'll start thinking about solving systems of linear equations on a computer. What can norms and condition numbers help us say about that?

Suppose we want to solve a linear system $Ax=b$ for x.

$$A\overset{\text{output}}{x} = \underset{\text{input}}{b}$$

$$A(x+\Delta x) = (b+\Delta b)$$

$\|\cdot\|$: some norm, doesn't matter which (but the condition number will depend on this choice.)

$$\text{Condition number} = \underbrace{\frac{\|\Delta x\|}{\|x\|}}_{\text{rel. error in output}} \Big/ \underbrace{\frac{\|\Delta b\|}{\|b\|}}_{\text{rel. error in input}} = \frac{\|\Delta x\| \cdot \|b\|}{\|x\| \cdot \|\Delta b\|}$$

$$= \frac{\|b\|}{\|x\|} \cdot \frac{\|\Delta x\|}{\|\Delta b\|}$$

$$= \frac{\|Ax\|}{\|x\|} \cdot \frac{\|\Delta x\|}{\|\Delta b\|}$$

$$= \frac{\|Ax\|}{\|x\|} \cdot \frac{\|A^{-1}\Delta b\|}{\|\Delta b\|}$$

$\|Ax\| \leq \|A\| \|x\|$

$$\leq \|A\| \cdot \|A^{-1}\|$$

Can you give an example?

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \rightsquigarrow \|A\|_2 = 3$$

$$A^{-1} = \begin{pmatrix} 1/3 & 0 \\ 0 & 1 \end{pmatrix} \rightsquigarrow \|A^{-1}\|_2 = 1$$

$$\rightsquigarrow \kappa_2 = \|A\|_2 \|A^{-1}\|_2 = 3$$

same!

Multiplying a vector by a matrix is an 'problem', too! So why so complicated?

What's the condition number of a 'matvec', a matrix-vector multiplication?

input, output

$$Ax = b$$

$$A(x+\Delta x) = (b+\Delta b)$$

Simplifying assumption: A is invertible. $\leadsto$ Let $B = A^{-1}$

output, input

$$Bb = x$$

That's linear system solving! We know the condition number of that!

$$\frac{\|\Delta b\|}{\|b\|} \le \|B\| \, \|B^{-1}\| \frac{\|\Delta x\|}{\|x\|}$$

$$\leadsto \frac{\|\Delta b\|}{\|b\|} \le \|A^{-1}\| \, \|A\| \frac{\|\Delta x\|}{\|x\|}$$

So what have we just learned?

The condition number of a matvec is *the same* as

that of solving a linear system.

Demo: Condition Number