5374 Fall '22

Numerical Linear Algebra Lecture 8

Today

* More on backward / forward error
* ~~Solvability of linear equations~~ next class
* Python : Working with sparse matrices

Mathematical problem

$$f : \underbrace{X}_{\text{(parameters of the problem / problem instances)}} \longrightarrow \underbrace{Y}_{\text{(solutions)}}$$

$$x \qquad\qquad f(x)$$

Algorithm: same as a Mathematical problem

$$\hat{f} : X \longrightarrow Y$$

except $\hat{f}$ we can implement in a computer.

The Problem: Given $f: X \to Y$, find an algorithm $\hat{f}: X \to Y$ such that $\hat{f}(x)$ is a good approximation to $f(x)$ for all $x \in X$.

Good approximation here means,

$$\| f(x) - \hat{f}(x) \| \quad \text{or} \quad \frac{\| f(x) - \hat{f}(x) \|}{\| f(x) \|}$$

is small.

To estimate this where we lack exact information of $f(x)$ we introduced the notion of backward error and forward error.

Forward error : The forward error (or relative forward error) made by an algorithm $\hat{f}$ when computing f at instance x, is

$$\| f(x) - \hat{f}(x) \| \quad \left(\text{or} \quad \frac{\| f(x) - f(\hat{x}) \|}{\| f(x) \|} \right)$$

Backward error : Suppose $\hat{f}(x)$ is not the solution to problem instance x, but rather the exact solution to a (hopefully) slightly different problem instance $\hat{x}$, i.e.

$$\exists\, \hat{x} \in X : \hat{f}(x) = f(\hat{x})$$

In this case, we look over all such $\hat{x}$'s and define the backward error of $\hat{f}$ at problem x, by

$$\inf \left\{ \|\hat{x} - x\| \;\middle|\; \hat{x} \text{ s.t. } f(\hat{x}) = f(x) \right\}$$

The relative backward error is defined similarly

$$\inf \left\{ \frac{\|\hat{x} - x\|}{\|x\|} \;\middle|\; \hat{x} \text{ s.t. } f(\hat{x}) = f(x) \right\}$$

Example (floating point arithmetic)
(AKA how machines round off computations)

In a subset of $\mathbb{R}$ denoted $\mathbb{F}$, we define operation $\oplus, \ominus, \otimes, \oslash$ that approximate the usual algebraic operations: Given $x, y \in \mathbb{F}$,

$$x \oplus y, \quad x \ominus y, \quad x \otimes y, \quad x \oslash y \;(\text{if } y \neq 0)$$

They have the following property

$$\frac{|x \oplus y - (x+y)|}{|x+y|} \leq \epsilon_m$$

$$\frac{|x \ominus y - (x-y)|}{|x-y|} \leq \epsilon_m$$

$$\frac{|x \otimes y - (x \times y)|}{|x \times y|} \leq \epsilon_m$$

$$\frac{|x \oslash y - x/y|}{|x/y|} \leq \epsilon_m$$

where ϵ_m is a small number associated to $\mathbb{F}$ called the machine epsilon.

Last but not least, for any $x \in \mathbb{R}$, $\exists$ $x^* \in \mathbb{F}$ such that $\frac{|x - x^*|}{|x|} \leq \epsilon_m$ (if $x = 0$, $|x^*| \leq \epsilon_m$)

With these definitions in hand, let us solve in $\mathbb{F}$, the equation

$$ax = b,$$

where $a, b \in \mathbb{R}$, $a \neq 0$

As a mathematical problem

$$X = \{ \overbrace{(a,b)}^{x} \mid a \in \mathbb{R}\setminus\{0\},\ b \in \mathbb{R} \}$$
$$Y = \mathbb{R}$$

$$f(a,b) := \{\text{the solution } x \text{ of } ax = b\}$$
$$= \frac{b}{a} .$$

$\hat{f}$ will be the algorithm that uses a computer to estimate f.

$$\hat{f}(a,b) = b^* \oslash a^*$$

(clearly this is an algorithm because we can instruct a computer to do it).

Observe: $*$ $|b^* - b| \le \epsilon_m |b|$, i.e.

$b^* = b(1+\varepsilon_1)$, where

$|\varepsilon_1| \le \epsilon_m$

$*$ Likewise,

$a^* = a(1+\varepsilon_2)$, where

$|\varepsilon_2| \le \epsilon_m$

* Then, $b^* \oplus a^*$ $= b(1+\varepsilon_1) \oplus a(1+\varepsilon_2)$

and

$$|b^* \oplus a^* - b^*/a^*| \le \epsilon_m \, |b^*/a^*|$$

i.e.

$$b^* \oplus a^* = \frac{b^*}{a^*}(1+\varepsilon_3)$$

where $|\varepsilon_3| \le \epsilon_m$.

* Combining all of the above, we have

$$b^* \oplus a^* = \frac{b(1+\varepsilon_1)(1+\varepsilon_3)}{a(1+\varepsilon_2)}$$

$$= \frac{b}{a} \cdot \left(\frac{(1+\varepsilon_1)(1+\varepsilon_3)}{(1+\varepsilon_2)} \right)$$

where $|\varepsilon_1|, |\varepsilon_2|, |\varepsilon_3| \le \epsilon_m$

Exercise: Show that

$$\frac{(1+\varepsilon_1)(1+\varepsilon_3)}{(1+\varepsilon_2)} = 1 + \varepsilon_4 + O(\epsilon_m^2)$$

where $\qquad |\varepsilon_4| \le 3\epsilon_m$

So $\qquad b^* \oplus\!\!\!\!\div\, a^* = \frac{b}{a}\left(1+\varepsilon_4 + O(\epsilon_m^2)\right)$

What is the forward error here?

$$f(a,b) - \hat{f}(a,b) = \frac{b}{a} - b^* \oplus\!\!\!\!\div\, a^*$$

$$= \frac{b}{a}\varepsilon_4 + O(\epsilon_m^2)$$

so $\qquad |f(a,b) - \hat{f}(a,b)| \le |f(a,b)|\, 3\epsilon_m + O(\epsilon_m^2)$

$$\frac{|f(a,b) - \hat{f}(a,b)|}{|f(a,b)|} \le 3\epsilon_m + O(\epsilon_m^2)$$

What about the backward error?

Well,

$$z := \hat{f}(a,b).$$

Then $\qquad z = \dfrac{b(1+\varepsilon_1)(1+\varepsilon_3)}{a(1+\varepsilon_2)}$

So

$$a(1+\varepsilon_2)z = b(1+\varepsilon_1)(1+\varepsilon_3)$$

If we define: $\hat{a} = a(1+\varepsilon_2)$

$\hat{b} = b(1+\varepsilon_1)(1+\varepsilon_3)$

Then $$\hat{a}z = \hat{b}$$

and we conclude that

$$\hat{f}(a,b) = f(\hat{a},\hat{b})$$

where $$\frac{\|\hat{a}-a\|}{\|a\|} + \frac{\|\hat{b}-b\|}{\|b\|} \leq |\varepsilon_2| + |(1+\varepsilon_1)(1+\varepsilon_3)-1| \leq 3\epsilon_m + \epsilon_m^2$$

This is the (relative) backward error of $\hat{f}$.

Next time: solvability of $Ax=b$, and a word about Lipschitz functions