Numerical stability

MATH 6610 Lecture 07

September 18, 2020

Trefethen & Bau: Lecture 14, 15

Numerical algorithms

Given $f : \mathbb{C}^n \to \mathbb{C}^m$, we wish to understand how roundoff errors affect evaluation of f.

This should depend on

- Loss of accuracy due to finite precision $\rightarrow$ due to floating point arithmetic
- Conditioning of f $\rightarrow$ sensitivity of f.

Numerical algorithms

Given $f : \mathbb{C}^n \to \mathbb{C}^m$, we wish to understand how roundoff errors affect evaluation of f.

This should depend on

- Loss of accuracy due to finite precision
- Conditioning of f

Let $\underbrace{\widetilde{f}}_{\text{algorithm}} : \mathbb{C}^n \to \mathbb{C}^m$ denote the actual algorithmic implementation of f.

We might hope that

exact output   algorithm output

$$\frac{\|f(x) - \widetilde{f}(x)\|}{\|f(x)\|} = \mathcal{O}(\epsilon_{\text{mach}}),$$

→ i.e. error depending on finite-precision rounding.

but this is typically too much to ask for if f is ill-conditioned.

factor of $\kappa_f(x)$
(relative cond. number of f)

Forward stability

We can, however, analyze the error in the algorithm $\widetilde{f}$.

??

Forward stability

We can, however, analyze the error in the algorithm $\widetilde{f}$. , $\widetilde{x}$: perturbation of x in floating point : $\frac{\|x-\widetilde{x}\|}{\|x\|} = \varepsilon_{mach}$

$- f(\widetilde{x}) + f(\widetilde{x})$

$$\frac{\|\widetilde{f}(x) - f(x)\|}{\|f(x)\|} \leqslant \frac{\|\widetilde{f}(x) - f(\widetilde{x})\|}{\|f(x)\|} + \underbrace{\frac{\|f(\widetilde{x}) - f(x)\|}{\|f(x)\|}}_{\text{independent of algorithm!}},$$

and this motivates a definition.

Recall : $\frac{\|f(\widetilde{x}) - f(x)\|}{\|f(x)\|} \leq \kappa_f(x) \cdot \frac{\|x-\widetilde{x}\|}{\|x\|}$ (def'n of κ_f).

$\leq \kappa_f(x) \cdot \varepsilon_{mach}$

Strategy: define "stability" as something that makes first term comparable to 2nd term (in magnitude)

Forward stability

We can, however, analyze the error in the algorithm $\widetilde{f}$.

$$\frac{\|\widetilde{f}(x) - f(x)\|}{\|f(x)\|} \leqslant \frac{\|\widetilde{f}(x) - f(\widetilde{x})\|}{\|f(x)\|} + \frac{\|f(\widetilde{x}) - f(x)\|}{\|f(x)\|},$$

and this motivates a definition.

Definition

An algorithm $\widetilde{f}$ is forward stable if, for every $x \in \mathbb{C}^n$, we have

algorithm output

"almost" the exact answer

$$\frac{\|\widetilde{f}(x) - f(\widetilde{x})\|}{\|f(\widetilde{x})\|} = \mathcal{O}(\epsilon_{\mathrm{mach}}),$$

for some $\widetilde{x}$ satisfying $\|x - \widetilde{x}\| = \|x\|\mathcal{O}(\epsilon_{\mathrm{mach}})$.

A forward stable algorithm gives an "approximately correct answer to a closely related question."

$f(\tilde{x}) \neq f(x)$

If $\tilde{f}$ is forward stable:

$$\frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} \leq \kappa_f(x) \cdot O(\varepsilon_{mach})$$

(from triangle inequality above)

Forward stability, II

$$\frac{\|\widetilde{f}(x) - f(\widetilde{x})\|}{\|f(\widetilde{x})\|} = \mathcal{O}(\epsilon_{\text{mach}}), \tag{1}$$

for some $\widetilde{x}$ satisfying $\|x - \widetilde{x}\| = \|x\|\mathcal{O}(\epsilon_{\text{mach}})$.

If $\widetilde{f}$ is forward stable, we can show that $\widetilde{f}$ produces a reasonable approximation to f.

(I.e. $\dfrac{\|f(x) - \widetilde{f}(x)\|}{\|f(x)\|} \leq \kappa_f \cdot \mathcal{O}(\varepsilon_{\text{mach}})$)

Forward stability, II

$$\frac{\|\widetilde{f}(x) - f(\widetilde{x})\|}{\|f(\widetilde{x})\|} = \mathcal{O}(\epsilon_{\mathrm{mach}}), \tag{1}$$

for some $\widetilde{x}$ satisfying $\|x - \widetilde{x}\| = \|x\|\mathcal{O}(\epsilon_{\mathrm{mach}})$.

If $\widetilde{f}$ is forward stable, we can show that $\widetilde{f}$ produces a reasonable approximation to f.

This error estimation procedure, requiring the establishment of (1), is *forward error analysis*.

Showing forward stability (1) is frequently technical and difficult.

Backward stability

A dual, but somewhat stranger+stronger notion of stability is backward stability.

Definition

An algorithm $\tilde{f}$ is backward stable if, for every $x \in \mathbb{C}^n$, we have

$$\tilde{f}(x) = f(\tilde{x}), \tag{2}$$

for some $\tilde{x}$ satisfying $\|x - \tilde{x}\| = \|x\| \mathcal{O}(\epsilon_{\text{mach}})$.

A backward stable algorithm gives an "exact answer to a closely related question."

1.) If $\tilde{f}$ is backward stable $\Longrightarrow$ $\underbrace{\dfrac{\|\tilde{f}(x) - f(\tilde{x})\|}{\|f(x)\|}}_{\text{first term in triangle inequality.}} = 0$

2.) Backward stability

$\Downarrow$

forward stability (def'n)

Backward stability

A dual, but somewhat stranger+stronger notion of stability is backward stability.

Definition

An algorithm $\widetilde{f}$ is backward stable if, for every $x \in \mathbb{C}^n$, we have

$$\widetilde{f}(x) = f(\widetilde{x}), \tag{2}$$

for some $\widetilde{x}$ satisfying $\|x - \widetilde{x}\| = \|x\| \mathcal{O}(\epsilon_{\mathrm{mach}})$.

A backward stable algorithm gives an "exact answer to a closely related question."

If $\widetilde{f}$ is backward stable, then showing accuracy of the algorithm $\widetilde{f}$ is easier.

Estimating error by establishing (2) is *backward error analysis*.

Backward stability is frequently easier to show than forward stability.

Examples

are actually true in IEEE 754.

Assume some floating-point axioms:

1. For each $x \in \mathbb{R}$, then $fl(x) = x(1+\epsilon)$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\text{mach}})$.
2. For floating-points numbers $x, y \in \mathbb{R}$, then $x \odot y = \underline{(x \cdot y)(1+\epsilon)}$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\text{mach}})$. $\quad (\cdot = +, -, \times)$

rounding error.

$x \mapsto fl(x) \Rightarrow$ floating point representation of x.

$\oplus, \ominus, \otimes \to$ floating-point implementations of $+, -, \times$.

Examples

Assume some floating-point axioms:

1. For each $x \in \mathbb{R}$, then $fl(x) = x(1+\epsilon)$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\mathrm{mach}})$.
2. For floating-points numbers $x, y \in \mathbb{R}$, then $x \odot y = (x \cdot y)(1+\epsilon)$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\mathrm{mach}})$. $\quad (\cdot = +, -, \times)$

Example

Given $x, y \in \mathbb{R}$, is the implementation $\widetilde{f}(x,y) := fl(x) \ominus fl(y)$ of $f(x,y) = x - y$ backward stable? (Yes)

(trying to show $\tilde{f}(x,y) = f(\tilde{x}, \tilde{y})$)

$$\tilde{f}(x,y) = fl(x) \ominus fl(y) = x(1+\epsilon_1) \ominus y(1+\epsilon_2),$$

axiom 1

$\epsilon_1, \epsilon_2 \leq O(\epsilon_{mach})$

$$= [x(1+\varepsilon_1) - y(1+\varepsilon_2)](1+\varepsilon_3), \quad \varepsilon_3 = O(\varepsilon_{mach}).$$

axiom 2

$$= x\underbrace{(1+\varepsilon_1)(1+\varepsilon_3)}_{1+O(\varepsilon_{mach})} - y\underbrace{(1+\varepsilon_2)(1+\varepsilon_3)}_{1+O(\varepsilon_{mach})}$$

$$= \underbrace{x(1+\varepsilon_4)}_{\tilde{x}} - \underbrace{y(1+\varepsilon_5)}_{\tilde{y}} = f(\tilde{x}, \tilde{y})$$

Note: f is <u>not</u> well-conditioned.

$$f(x,y) = x - y \longrightarrow \|J\|_2 = \sqrt{2}$$

$$\kappa_f(x,y) = \frac{\|J\|_2 \, \|(x,y)\|_2}{\|f(x,y)\|_2} = \frac{\sqrt{2}\sqrt{x^2+y^2}}{|x-y|}$$

can be big if $|x-y|$ is small.

Examples

Assume some floating-point axioms:

1. For each $x \in \mathbb{R}$, then $fl(x) = x(1+\epsilon)$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\text{mach}})$.
2. For floating-points numbers $x, y \in \mathbb{R}$, then $x \odot y = (x \cdot y)(1+\epsilon)$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\text{mach}})$. $\quad (\cdot = +, -, \times)$

Example

Given $x, y \in \mathbb{R}$, is the implementation $\widetilde{f}(x,y) := fl(x) \ominus fl(y)$ of $f(x,y) = x - y$ backward stable?

Example

Given $x \in \mathbb{R}$, is the implementation $\widetilde{f}(x) := 1 \oplus fl(x)$ of $f(x) = 1 + x$ backward stable? (No)

(note: $1 = fl(1)$)

$\tilde{f}(x,y) := fl(x) \oplus fl(y)$

Is backward stable

Examples

Assume some floating-point axioms:

1. For each $x \in \mathbb{R}$, then $fl(x) = x(1+\epsilon)$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\text{mach}})$.
2. For floating-points numbers $x, y \in \mathbb{R}$, then $x \odot y = (x \cdot y)(1+\epsilon)$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\text{mach}})$. $\quad (\cdot = +, -, \times)$

Example

Given $x, y \in \mathbb{R}$, is the implementation $\widetilde{f}(x,y) := fl(x) \ominus fl(y)$ of $f(x,y) = x - y$ backward stable?

Example

Given $x \in \mathbb{R}$, is the implementation $\widetilde{f}(x) := 1 \oplus fl(x)$ of $f(x) = 1 + x$ backward stable?

Example

Given $x, y \in \mathbb{R}^n$, the floating-point implementation of $f(x,y) = y^T x$ is backward stable. ✓

Examples

Assume some floating-point axioms:

1. For each $x \in \mathbb{R}$, then $fl(x) = x(1+\epsilon)$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\text{mach}})$.
2. For floating-points numbers $x, y \in \mathbb{R}$, then $x \odot y = (x \cdot y)(1+\epsilon)$ for $|\epsilon| \leqslant \mathcal{O}(\epsilon_{\text{mach}})$. $\quad (\cdot = +, -, \times)$

Example

Given $x, y \in \mathbb{R}$, is the implementation $\widetilde{f}(x,y) := fl(x) \ominus fl(y)$ of $f(x,y) = x - y$ backward stable?

Example

Given $x \in \mathbb{R}$, is the implementation $\widetilde{f}(x) := 1 \oplus fl(x)$ of $f(x) = 1 + x$ backward stable?

Example

Given $x, y \in \mathbb{R}^n$, the floating-point implementation of $f(x,y) = y^T x$ is backward stable.

Example

The floating-point implementation of $f(x,y) = xy^T$ is *not* backward stable.

Issue: F for outer products can produce answers that are of rank greater than 1.