Technische Universität Braunschweig

Institute for Numerical Analysis

Ramp Up Mathematics — Numerical Analysis Ramp Up for Data Science

Matthias Bollhöfer, SS 2024

Contents I

Contents II

Reviewing and Comparing Numerical Methods

1. *Efficiency*: computational amount of work, memory consumption, parallelization
2. *Reliability and fault resilience*: how do failures of a parts of the method influence the overall method (e.g., poor approximation, approximation failures, ...)
3. *Accuracy*: *all possible* errors and their influence on the computed result
   - *Round off errors*: $\surd$
   - *Method based errors*: depends on the method (e.g., linear approximation of a nonlinear function)
   - *Data errors*: input data typically inaccurate.
     Solving a problem refers to evaluation of a function $F : D \longrightarrow W$.
     Instead of an exact value x we have a perturbed $\tilde{x}$.
     resulting output error in when evaluation the function: $F(\tilde{x}) - F(x)$.

Conditioning

perturbed input data: $\tilde{x}_1 = x_1 + \Delta x_1, \ \ldots, \ \tilde{x}_n = x_n + \Delta x_n,$

Determine output error $F(\tilde{x}) - F(x)$ using Taylor expansion:

$$\begin{aligned} F(\tilde{x}_1, \ldots, \tilde{x}_n) \quad &= \quad F(x_1, \ldots, x_n) + \frac{\partial F(x_1, \ldots, x_n)}{\partial x_1} \Delta x_1 + \cdots + \frac{\partial F(x_1, \ldots, x_n)}{\partial x_n} \Delta x_n \\ & \quad + \text{higher order terms (h.o.t.)} \end{aligned}$$

$$\begin{aligned} \Rightarrow \underbrace{\frac{|F(\tilde{x}_1, \ldots, \tilde{x}_n) - F(x_1, \ldots, x_n)|}{|F(x_1, \ldots, x_n)|}}_{\text{rel. output error}} \quad &\leqslant \quad \frac{\left|\frac{\partial F(x_1,\ldots,x_n)}{\partial x_1} \Delta x_1\right| + \cdots + \left|\frac{\partial F(x_1,\ldots,x_n)}{\partial x_n} \Delta x_n\right|}{|F(x_1, \ldots, x_n)|} + \text{h.o.t.} \\ &\leqslant \quad \max_{i=1,\ldots,n} |\Delta x_i| \cdot \frac{\sum_{i=1}^n \left|\frac{\partial F(x_1,\ldots,x_n)}{\partial x_i}\right|}{|F(x_1, \ldots, x_n)|} + \text{h.o.t.} \\ &= \quad \underbrace{\frac{\|\Delta x\|_\infty}{\|x\|_\infty}}_{\text{rel. input error}} \cdot \underbrace{\frac{\|x\|_\infty \sum_{i=1}^n \left|\frac{\partial F(x_1,\ldots,x_n)}{\partial x_i}\right|}{|F(x_1, \ldots, x_n)|}}_{\text{amplification factor}} + \text{h.o.t.} \end{aligned}$$

Conditioning

We conclude for a differentiable function $F : D \subset \mathbb{R}^n \to W \subset \mathbb{R}^m$ that

$$\frac{\|\vec{F}(\tilde{\vec{x}}) - \vec{F}(\vec{x})\|}{\|\vec{F}(\vec{x})\|} \leqslant \frac{\|\frac{\partial \vec{F}(\vec{x})}{\partial \vec{x}}\| \cdot \|\vec{x}\|}{\|\vec{F}(\vec{x})\|} \cdot \frac{\|\tilde{\vec{x}} - \vec{x}\|}{\|\vec{x}\|} + \text{h.o.t.}$$

relative output error $\frac{\|\vec{F}(\tilde{\vec{x}}) - \vec{F}(\vec{x})\|}{\|\vec{F}(\vec{x})\|}$ essentially described by the size of the derivative
Suppose that

$$\frac{\|\frac{\partial \vec{F}(\vec{x})}{\partial \vec{x}}\| \cdot \|\vec{x}\|}{\|\vec{F}(\vec{x})\|} \leqslant C \text{ for all } \vec{x},$$

then we obtain

$$\frac{\|\vec{F}(\tilde{\vec{x}}) - \vec{F}(\vec{x})\|}{\|\vec{F}(\vec{x})\|} \lessapprox C \cdot \frac{\|\tilde{\vec{x}} - \vec{x}\|}{\|\vec{x}\|}$$

Conclusion: In this case we conclude that $\vec{F}(\vec{x})$ is well conditioned

Condition

Definition 1.1

Given a function $F : D \subset \mathbb{R}^n \to W \subset \mathbb{R}^m$ we define $\kappa(\vec{F})$ as smallest possible number such that

$$\frac{\|\vec{F}(\tilde{\vec{x}}) - \vec{F}(\vec{x})\|}{\|\vec{F}(\vec{x})\|} \lessapprox \kappa(\vec{F}) \cdot \frac{\|\tilde{\vec{x}} - \vec{x}\|}{\|\vec{x}\|}$$

Suppose that F is differentiable. If its partial derivative satisfies $\frac{\|\frac{\partial \vec{F}(\vec{x})}{\partial \vec{x}}\| \cdot \|\vec{x}\|}{\|\vec{F}(\vec{x})\|} \leqslant C$, then it follows that $\kappa(\vec{F}) \leqslant C$

Example 1.1 (Sum)

$$F(x, y) = x + y \Rightarrow \frac{|F(\tilde{x}, \tilde{y}) - F(x, y)|}{|F(x, y)|} = \qquad\qquad \leqslant$$

Error Analysis of Numerical Methods

- Investigate the implementation of our algorithm on a computer
- Instead of a function $F(x_1, \ldots, x_n)$ our implementation on a computer yields $\tilde{F}(x_1, \ldots, x_n)$ including round off errors

1. *Forward analysis*: track errors step by step, determine bounds for parts of the computation. What we get at the end is

$$\tilde{F}(x_1, \ldots, x_n) = F(x_1, \ldots, x_n) + \delta,$$

   where δ refers to the accumulated error.

2. *Backward analysis*: rewrite the accumulated error as *exact evaluation of perturbed input data*, i.e.,

$$\tilde{F}(x_1, \ldots, x_n) = F(\tilde{x}_1, \ldots, \tilde{x}_n)$$

   with perturbed input data $\tilde{x}_1, \ \ldots, \ \tilde{x}_n$.

well-conditioned problem plus stable algorithm $\Longrightarrow$ reliable result

Error Analysis

Example 1.2

Consider

$$F(x_1,x_2)=x_1+x_2,\ \tilde{F}(x_1,x_2)=x_1\oplus x_2.$$

forward analysis

$$x_1\oplus x_2 = \underbrace{x_1+x_2}_{F(x_1,x_2)}+\underbrace{\varepsilon(x_1+x_2)}_{\delta} \qquad |\varepsilon|\leqslant \epsilon,$$

$$|\delta|\leqslant \epsilon(|x_1|+|x_2|).$$

backward analysis

$$x_1\oplus x_2 = \underbrace{x_1(1+\varepsilon)+x_2(1+\varepsilon)}_{F(\tilde{x}_1,\tilde{x}_2)} \qquad |\varepsilon|\leqslant \epsilon.$$