9. Conditioning and stability

Last time

- Rules for derivatives via new algebra
- Dual numbers
- Implementation
- Directional derivatives, partial derivatives, jacobians

Goals for today

- Conditioning: Sensitivity of problems
- Catastrophic cancellation

Perturbing input data

- Suppose have (numerical) problem given by function ϕ:

Perturbing input data

- Suppose have (numerical) problem given by function ϕ:
- Given input x, calculate output $y = \phi(x)$

Perturbing input data

- Suppose have (numerical) problem given by function ϕ:
- Given input x, calculate output $y = \phi(x)$
- How **sensitive** is output to input?

Perturbing input data

- Suppose have (numerical) problem given by function ϕ:
- Given input x, calculate output $y = \phi(x)$
- How **sensitive** is output to input?

- Suppose perturb input by Δx.
- How large is resulting effect Δy on output?

Perturbing input data

- Suppose have (numerical) problem given by function ϕ:
- Given input x, calculate output $y = \phi(x)$
- How **sensitive** is output to input?

- Suppose perturb input by Δx.
- How large is resulting effect Δy on output?

- Example: Intersection of two thick lines at angle

Absolute errors

- Suppose $\widehat{x}$ is approximation of input x
- Define $\Delta x := \widehat{x} - x$ as error on the input

Absolute errors

- Suppose $\widehat{x}$ is approximation of input x
- Define $\Delta x := \widehat{x} - x$ as error on the input

- Calculated output is $\widehat{y} := \phi(\widehat{x}) = \phi(x + \Delta x)$

Absolute errors

- Suppose $\hat{x}$ is approximation of input x
- Define $\Delta x := \hat{x} - x$ as error on the input

- Calculated output is $\hat{y} := \phi(\hat{x}) = \phi(x + \Delta x)$

- Δx and Δy are **absolute errors**

Relative errors

- Size of absolute error Δy depends on size of input

Relative errors

- Size of absolute error Δy depends on size of input
- Usually more interested in **relative errors**

Relative errors

- Size of absolute error Δy depends on size of input
- Usually more interested in **relative errors**
- Divide absolute error by true value:

$$\delta x := \frac{\Delta x}{x} = \frac{\hat{x} - x}{x}; \quad \delta y := \frac{\Delta y}{y} = \frac{\hat{y} - y}{y}$$

Relative errors

- Size of absolute error Δy depends on size of input
- Usually more interested in **relative errors**
- Divide absolute error by true value:

$$\delta x := \frac{\Delta x}{x} = \frac{\hat{x} - x}{x}; \quad \delta y := \frac{\Delta y}{y} = \frac{\hat{y} - y}{y}$$

- So $\hat{x} = x + \Delta x = x(1 + \delta x)$

Significant / accurate digits

- Suppose have approximation $\hat{x}$ of true value x
- How **accurate** is it? Relative error is $|\delta x| = \left|\frac{\hat{x}-x}{x}\right|$

Significant / accurate digits

- Suppose have approximation $\hat{x}$ of true value x
- How **accurate** is it? Relative error is $|\delta x| = \left|\frac{\hat{x}-x}{x}\right|$

- Express number of **significant** or **accurate digits**:

$$d = -\log_{10} \left| \frac{\hat{x} - x}{x} \right|$$

- Gives number of digits after which $\hat{x}$ and x differ

Conditioning

- For some problems, perturbing input does not affect output too much: **well-conditioned**
- For others, perturbing input can change output drastically: **ill-conditioned**

Conditioning

- For some problems, perturbing input does not affect output too much: **well-conditioned**
- For others, perturbing input can change output drastically: **ill-conditioned**

- How can we **measure** conditioning?

Absolute condition number

- Given problem ϕ at input x
- Define **absolute condition number** $\widehat{\kappa}_{\phi}(x)$:

Absolute condition number

- Given problem ϕ at input x
- Define **absolute condition number** $\hat{\kappa}_\phi(x)$:

$$\hat{\kappa}_\phi(x) = \frac{\|\Delta y\|}{\|\Delta x\|}$$

Absolute condition number

- Given problem ϕ at input x
- Define **absolute condition number** $\hat{\kappa}_\phi(x)$:

$$\hat{\kappa}_\phi(x) = \frac{\|\Delta y\|}{\|\Delta x\|}$$

- $\| \cdot \|$ are suitable **norms** measuring length of vectors
- For real numbers, just take absolute value

Absolute condition number

- Given problem ϕ at input x
- Define **absolute condition number** $\hat{\kappa}_\phi(x)$:

$$\hat{\kappa}_\phi(x) = \frac{\|\Delta y\|}{\|\Delta x\|}$$

- $\| \cdot \|$ are suitable **norms** measuring length of vectors
- For real numbers, just take absolute value

- $\hat{\kappa}_\phi(x, \epsilon)$: take maximum over inputs $\|\Delta x\| < \epsilon$

Absolute condition number II

- What happens if input $|\Delta x| \to 0$?

Absolute condition number II

- What happens if input $|\Delta x| \to 0$?
- Have $\Delta y = \phi(x + \Delta x) - \phi(x)$

Absolute condition number II

- What happens if input $|\Delta x| \to 0$?
- Have $\Delta y = \phi(x + \Delta x) - \phi(x)$
- So $\Delta y = \phi'(x)\, \Delta x$ to first order

Absolute condition number II

- What happens if input $|\Delta x| \to 0$?
- Have $\Delta y = \phi(x + \Delta x) - \phi(x)$
- So $\Delta y = \phi'(x)\,\Delta x$ to first order

- Hence $\hat{\kappa}_\phi(x) = |\phi'(x)|$

Relative condition number

- Define **relative condition number** $\kappa_\phi(x)$ by

$$\kappa_\phi(x) = \frac{\|\delta y\|}{\|\delta x\|}$$

- Here we have the *relative* errors on top and bottom

Relative condition number

- Define **relative condition number** $\kappa_\phi(x)$ by

$$\kappa_\phi(x) = \frac{\|\delta y\|}{\|\delta x\|}$$

- Here we have the *relative* errors on top and bottom

- Taking limit as $|\Delta x| \to 0$, obtain

$$\kappa_\phi(x) = \lim_{\Delta x \to 0} \left| \frac{\frac{\phi(x+\Delta x) - \phi(x)}{\phi(x)}}{\frac{\Delta x}{x}} \right| = \left| \frac{x\, \phi'(x)}{\phi(x)} \right|$$

Relative condition number

- Define **relative condition number** $\kappa_\phi(x)$ by

$$\kappa_\phi(x) = \frac{\|\delta y\|}{\|\delta x\|}$$

- Here we have the *relative* errors on top and bottom

- Taking limit as $|\Delta x| \to 0$, obtain

$$\kappa_\phi(x) = \lim_{\Delta x \to 0} \left| \frac{\frac{\phi(x+\Delta x)-\phi(x)}{\phi(x)}}{\frac{\Delta x}{x}} \right| = \left| \frac{x\,\phi'(x)}{\phi(x)} \right|$$

Example: Condition number of addition

- Addition of two numbers is a very basic operation
- We hope that it is well-conditioned. Is it?

Example: Condition number of addition

- Addition of two numbers is a very basic operation
- We hope that it is well-conditioned. Is it?
- Take problem $\phi(x) = x + a$ with a fixed a
- $\Delta y = \phi(x + \Delta x) - \phi(x)$
- $= a + (x + \Delta x) - (a + x) = \Delta x$
- Abs. condition number $\hat{\kappa}_\phi = |\phi'(x)| = 1$ – well-behaved
- But we really care about *relative* condition number

Condition number of addition II

- Problem $\phi(x) = a + x$ with fixed a
- But relative condition number is

$$\kappa_\phi = \left| \frac{x\,\phi'(x)}{\phi(x)} \right| = \left| \frac{x}{a+x} \right|$$

- Note: $\phi(x)$ appears in *denominator*
- So κ is *large* when $x \simeq -a$
- **Catastrophic cancellation**: loss of digits of accuracy when subtracting two numbers that are close together
- Common task in numerical analysis: **identify + eliminate** catastrophic cancellation

Revisiting quadratic equations

- Let's revisit the topic of solving quadratic equations

$$f(x) = ax^2 + bx + c = 0$$

- Problem ϕ: inputs: (a, b, c); outputs: roots $x_\pm$

$$x_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Revisiting quadratic equations

- Let's revisit the topic of solving quadratic equations

$$f(x) = ax^2 + bx + c = 0$$

- Problem ϕ: inputs: (a, b, c); outputs: roots $x_{\pm}$

$$x_{\pm} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

- What can go wrong?

Revisiting quadratic equations

- Let's revisit the topic of solving quadratic equations

$$f(x) = ax^2 + bx + c = 0$$

- Problem ϕ: inputs: (a, b, c); outputs: roots $x_\pm$

$$x_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

- What can go wrong?

- Intuitively: if $b > 0$ and ac is small, then x_+ suffers from catastrophic cancellation

Revisiting quadratic equations

- Let's revisit the topic of solving quadratic equations

$$f(x) = ax^2 + bx + c = 0$$

- Problem ϕ: inputs: (a, b, c); outputs: roots $x_\pm$

$$x_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

- What can go wrong?

- Intuitively: if $b > 0$ and ac is small, then x_+ suffers from catastrophic cancellation

Eliminating cancellation for quadratics

- The root x_- is unaffected by catastrophic cancellation (sum of two quantities of similar size)
- Can we calculate x_+ from x_-?

Eliminating cancellation for quadratics

- The root $x_\_$ is unaffected by catastrophic cancellation (sum of two quantities of similar size)
- Can we calculate x_+ from $x_\_$?
- Factor quadratic as $f(x) = a(x - x_\_)(x - x_+)$
- Then $a\, x_\_\, x_+ = c$, so $x_+ = \frac{c}{a\, x_\_}$

Condition number of quadratic equations

- How calculate condition number of root of quadratic equation?

Condition number of quadratic equations

- How calculate condition number of root of quadratic equation?
- Suppose move a to $a + \Delta a$
- Root x_+ moves to $x_+ + \Delta x_+$

Condition number of quadratic equations

- How calculate condition number of root of quadratic equation?
- Suppose move a to $a + \Delta a$
- Root x_+ moves to $x_+ + \Delta x_+$
- $\kappa = \left| \frac{x_+}{a} \frac{\partial x_+}{\partial a} \right|$
- Find $\kappa = \left| \frac{x_+}{x_+ - x_-} \right|$
- So condition number is large when roots are close
- Visualize by perturbing function near double root

Stability of algorithms

- Notion of “condition number” refers to a **problem**
- Solving a problem on the computer needs an **algorithm**
- Conditioning is independent of the algorithm

Stability of algorithms

- Notion of “condition number” refers to a **problem**
- Solving a problem on the computer needs an **algorithm**
- Conditioning is independent of the algorithm

- An algorithm replaces problem ϕ with alternative $\widehat{\phi}$
- In general the result cannot be better than the condition number suggests
- If algorithm is much **worse** than expected from conditioning, algorithm is **unstable**; otherwise **stable**

Backward error

- A very useful point of view is that of **backward error**:
- Suppose algorithm $\hat{\phi}$ approximates a problem ϕ
- Then $\hat{y} := \hat{\phi}(x)$ approximates exact result $y := \phi(x)$

Backward error

- A very useful point of view is that of **backward error**:
- Suppose algorithm $\widehat{\phi}$ approximates a problem ϕ
- Then $\hat{y} := \widehat{\phi}(x)$ approximates exact result $y := \phi(x)$
- So far have studied **forward error** $\Delta y := \hat{y} - y$

Backward error

- A very useful point of view is that of **backward error**:
- Suppose algorithm $\hat{\phi}$ approximates a problem ϕ
- Then $\hat{y} := \hat{\phi}(x)$ approximates exact result $y := \phi(x)$

- So far have studied **forward error** $\Delta y := \hat{y} - y$
- Instead: “calculated result = exact solution for nearby input?”
- i.e. want **backward error** Δx such that

$$\hat{y} = f(x + \Delta x)$$

Summary

- (Absolute and relative) condition number for a problem
- Catastrophic cancellation in subtraction and its avoidance
- Stability
- Backward error