Lesson 12: Conditioning of the Problem

Notation:

$$\begin{cases} \text{Problem: } f \qquad \text{Data: } x \\ \text{Problem Answer: } f(x) \end{cases}$$

$$\begin{cases} \text{Problem: } f \qquad \text{Perturbed Data: } x + \Delta x \\ \text{Perturbed Problem Answer: } f(x+\Delta x) \end{cases}$$

Condition # of Problem: (Absolute) – The absolute change in the problem answer divided by the absolute change in the problem data (as the data change $\to 0$)

$$\hat{k} = \lim_{\delta \to 0} \sup_{\|\Delta x\| \le \delta} \frac{\| f(x+\Delta x) - f(x) \|}{\| \Delta x \|}$$

Note: x can be a vector, so f is a function of several variables. $f(x)$ could also be a vector: $f_1(x_1, x_2, \ldots, x_n)$, $f_2(x_1, x_2, \ldots, x_n)$, etc.

$$\hat{k} = \| J(x) \|, \quad J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \\ \vdots & \vdots & & \vdots \end{pmatrix}$$

if f is differentiable.

$J(x)$ is the Jacobian matrix.

Example: $f(x) = x_1 - x_2$. Find $\hat{\kappa}$.

$$J = \left(\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \right) = \begin{bmatrix} 1 & -1 \end{bmatrix}$$

$$\Delta x = \begin{pmatrix} \cancel{x_1} + \Delta x_1 \\ \cancel{x_2} + \Delta x_2 \end{pmatrix}$$

In ∞-norm, $\|J\|_\infty = 2$

$$\hat{\kappa} = \lim_{\delta \to 0} \sup_{\|\Delta x\| \le \delta} \left\| \left(\frac{x_1 + \Delta x_1 - (x_2 + \Delta x_2) - (x_1 - x_2)}{\|\Delta x\|} \right) \right\|$$

$$= \lim_{\delta \to 0} \sup_{\|\Delta x\| \le \delta} \frac{\|\Delta x_1 - \Delta x_2\|}{\left\| \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} \right\|} = 2 \; !$$

(The max occurs when Δx_2 + Δx_1 have opposite signs.)

$\|\Delta x_1 - \Delta x_2\|_\infty \le \|\Delta x_1\|_\infty + \|\Delta x_2\|_\infty \le 2 \max\{\|x_1\|_\infty, \|x_2\|_\infty\}$

$\left\| \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} \right\|_\infty = \max\{\|x_1\|_\infty, \|x_2\|_\infty\}$.

<u>Condition # of Problem</u> (Relative) – The relative change in problem answer divided by the relative change in the problem data (as the data change $\to 0$)

$$\kappa = \lim_{\delta \to 0} \sup_{\|\Delta x\| \le \delta} \frac{\dfrac{\|f(x+\Delta x) - f(x)\|}{\|f(x)\|}}{\dfrac{\|\Delta x\|}{\|x\|}}$$

For numerical computational purposes, we are more interested in whether the problem has a large relative condition number κ.

Example: For problem $f(x) = x_1 - x_2$. Find κ.

$$\kappa = \lim_{\delta \to 0} \sup_{\|\Delta x\| \le \delta} \frac{\dfrac{\|\Delta x_1 - \Delta x_2\|}{\|x_1 - x_2\|}}{\dfrac{\left\|\begin{pmatrix}\Delta x_1 \\ \Delta x_2\end{pmatrix}\right\|}{\left\|\begin{pmatrix}x_1 \\ x_2\end{pmatrix}\right\|}}$$

$$\|\Delta x_1 - \Delta x_2\|_\infty \le \|\Delta x_1\|_\infty + \|\Delta x_2\|_\infty \le 2 \max\{|\Delta x_1|, |\Delta x_2|\}.$$

$$\|x_1 - x_2\|_\infty = |x_1 - x_2| \quad ; \quad \|x\|_\infty = \max\{|x_1|, |x_2|\}$$

So, $\underbrace{\kappa = \frac{2}{|x_1 - x_2|} \cdot \max\{|x_1|, |x_2|\}}$. As $x_1 - x_2$ becomes small, this problem is poorly conditioned in the relative sense. We have seen this before as catastrophic cancellation.

<u>Example</u> : Finding roots of polynomials is an ill-conditioned problem.

Let $p(x) = \prod_{i=1}^{20}(x-i)$ (Has roots $1, 2, 3, \ldots, 20$)

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{19} x^{19} + x^{20}$$

$$a_{15} \approx 1.67 \times 10^9$$

Suppose we perturb the data (a_{15}) to be $a_{15} + \delta a_{15}$ and find the roots of the perturbed problem

$$a_0 + a_1 x + \cdots + (a_{15} + \delta a_{15}) x^{15} + \cdots + a_{19} x^{19} + x^{20}.$$

and ask how the 15th root of this problem changes due to the perturbation of a_{15}, we get

$$K = \frac{\dfrac{|\delta x_{15}|}{|x_{15}|}}{\dfrac{|\delta a_{15}|}{|a_{15}|}} \approx 5.1 \times 10^{13}$$

(see page 92, book)
(See picture, page 93)

<u>Note</u>: This is why we do not try to find eigenvalues of matrices by finding the roots of the characteristic polynomial given by $\det|A - \lambda I| = 0$!

Often, algorithms we develop to solve mathematical problems using linear algebra require matrix multiplication (Ax). How well-conditioned in the relative sense is this problem?

Problem Answer: Ax Data: x

Find κ relative to perturbations in x:

$$\kappa = \sup_{\Delta x} \frac{\dfrac{\| A(x+\Delta x) - Ax \|}{\|Ax\|}}{\dfrac{\| x + \Delta x - x \|}{\|x\|}} = \sup_{\Delta x} \frac{\| A \cdot \Delta x \|}{\dfrac{\|\Delta x\| \cdot \|Ax\|}{\|x\|}}$$

$$\kappa = \frac{\|x\|}{\|Ax\|} \cdot \frac{\| A \cdot \Delta x \|}{\|\Delta x\|}$$. Note $\dfrac{\|A \cdot \Delta x\|}{\|\Delta x\|} \leq \|A\|$!

So $\kappa \leq \dfrac{\|x\|}{\|Ax\|} \cdot \|A\|$. Note $\dfrac{\|x\|}{\|Ax\|} = \dfrac{\|A^{-1} y\|}{\|y\|} \leq \|A^{-1}\|$ (if A nonsingular)

So $\kappa \leq \|A^{-1}\| \cdot \|A\|$ (A, nonsingular)

$\kappa \leq K(A)$ where $K(A) = \|A^{-1}\| \cdot \|A\|$

What about the problem of solving $Ax = b$ (A, nonsingular) due to perturbations in data b? This is the same as

$$x = A^{-1} \cdot b$$

the <u>matrix multiplication</u> problem with matrix (A^{-1}) and data (b) that we already solved so,

$$\kappa \leq \kappa(A^{-1}) = \underline{\kappa(A)}, \text{ same as before.}$$

<u>Note:</u> $\kappa = \sup_{\Delta b} \dfrac{\dfrac{\| A^{-1}(b+\Delta b) - A^{-1}b \|}{\| A^{-1} b \|}}{\dfrac{\|\Delta b\|}{\|b\|}} \leq \kappa(A^{-1}) = \kappa(A)$!

Of interest to us when we study the solution of $Ax = b$ using the algorithm of Gaussian Elimination, is how well conditioned (or ill-conditioned) is the <u>problem</u>: find $x \ni Ax = b$ relative to perturbations of A.

So problem answer $= x$ and problem data $= A$. Let E be the perturbation of matrix A, and call $x + \Delta x$ the solution to the problem

$$\underbrace{(A+E)}_{\text{perturbed } A}\underbrace{(x+\Delta x)}_{\text{perturbed solution}} = \underbrace{b}_{\text{didn't perturb}} \qquad (1)$$

$$R = \frac{\text{relative error in answer}}{\text{relative error in data}} \text{ as perturbation} \to 0.$$

$$R = \sup_{\|E\|} \frac{\dfrac{\|x+\Delta x - x\|}{\|x\|}}{\dfrac{\|E\|}{\|A\|}} = \sup_{\|E\|} \frac{\dfrac{\|\Delta x\| \cdot \|A\|}{\|x\|}}{\|E\|} = \sup_{\|E\|} \frac{\|\Delta x\|\,\|A\|}{\|E\| \cdot \|x\|}$$

Use (1) to find Δx in terms of other terms.

$$\cancel{Ax} + A\Delta x + Ex + E\Delta x = \cancel{b}$$

$$A\Delta x = -Ex - \underbrace{E\Delta x}_{\text{doubly infinitesimal drop out.}}$$

$$\Delta x \simeq -A^{-1}Ex$$

So, $\|\Delta x\| \le \|A^{-1}\| \cdot \|E\| \cdot \|x\|$. Plug this in to get

$$K = \sup_{\|E\|} \frac{\|\Delta x\| \cdot \|A\|}{\|E\| \cdot \|x\|} \le \sup_{\|E\|} \frac{\{\|A^{-1}\| \cdot \cancel{\|E\|} \cdot \cancel{\|x\|}\} \cdot \|A\|}{\cancel{\|E\|} \cdot \cancel{\|x\|}}$$

$$= \|A^{-1}\| \cdot \|A\|.$$

Or $K \le \|A^{-1}\| \cdot \|A\|$. Further investigation (page 95) shows this bd is sharp and indeed $K = \|A^{-1}\| \cdot \|A\| = K(A)$.

Lesson 13. Floating Point Arithmetic

On the computer, we can only represent exactly the <u>floating point numbers</u>

$$a \times 10^{b} \quad \text{(base 10)}$$

$a =$ mantissa

$b =$ exponent

$a = x.xxx$

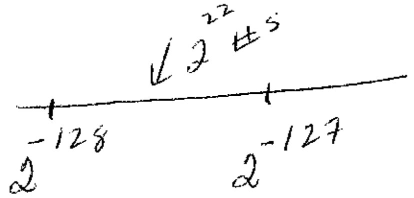
↑ nonzero

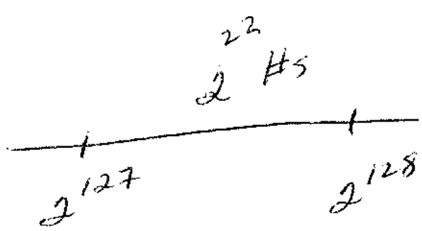
- largest floating point number (single precision)

$$\underbrace{(1.1\ldots1_2)}_{\text{24 bits for mantissa including sign}} \times 2^{\overset{\uparrow\ \text{8 bits for exponent including sign.}}{127}} \approx 10^{38}$$

- smallest $\approx 10^{-38}$ (single precision)

- 2^{22} #s (floating) between each power of 2 !





↑ floating pts #s clustered around zero

CHOP : 1.234|56... ON A 4-digit machine (base 10).

ROUND: 1.235 round up left-off part

MACHINE EPSILON ; (ϵ_{mach}) – Typically defined as the smallest # that can be added to 1 that gives a result bigger than 1.

$$1 + \epsilon_{machine} > 1$$

Adding Small #s to Large #s:

$$x + y = x\left(1 + \frac{y}{x}\right) . \text{ If } \left\|\frac{y}{x}\right\| < \epsilon_{mach},$$

$$\text{then } x + y = x(1+0) = X \quad !$$

Floating Pt. Representation of x :

$$fl(x) = x(1+\delta x) , \quad |\delta x| \le \epsilon_{mach}$$

$$\Rightarrow \quad fl(x) - x = x\,\delta x \quad \Rightarrow \quad |fl(x) - x| = |x \cdot \delta x|$$

$$\Rightarrow \quad \frac{|fl(x) - x|}{|x|} = |\delta x| \le \epsilon_{mach} .$$

"ϵ_{mach} is the max possible relative error in representing x."

Floating Point Arithmetic : Let $\circledast$ stand for floating pt. add, sub, mult, or divide. Then

$$x \circledast y = (x * y)(1+\epsilon)$$

$$|\epsilon| \leq \epsilon_{mach}$$

$$\Rightarrow \frac{|(x \circledast y) - (x * y)|}{|x * y|} \leq \epsilon_{mach}$$

"Every operation of floating pt. arithmetic is exact up to a relative error of size at most ϵ_{mach}".

Lesson 14: Stability of the Algorithm

In Lesson 12, we were interested in the conditioning of the problem. Even the best algorithms can not be expected to perform well on an ill-conditioned problem.

So what makes an algorithm a good one?

It is unfair to say an algorithm is good only if it produces accurate answers!

Accuracy: Notation: f is the problem
x is the data
$\tilde{f}$ is the algorithm
$f(x)$ is the problem's answer
$\tilde{f}(x)$ is the answer the algorithm gives if data x is input.

Relative Error in the computation $= \dfrac{\| \tilde{f}(x) - f(x) \|}{\| f(x) \|}$

Absolute Error in the computation $= \| \tilde{f}(x) - f(x) \|$.

If Relative error $= O(\epsilon_{mach})$, then we say the computation is accurate.

Accuracy is not a good measure of algorithm behavior, since the algorithm could be tossed for performing badly on an ill-conditioned problem. What can we reasonably expect then?

$\boxed{\text{Stability}}$: An algorithm $\tilde{f}$ for a problem f is stable if for each data $x \in X$,

$$\frac{\|\tilde{f}(x) - f(\tilde{x})\|}{\|f(\tilde{x})\|} = O(\epsilon_{mach})$$

for some $\tilde{x}$, $\dfrac{\|\tilde{x} - x\|}{\|x\|} = O(\epsilon_{mach})$

Note: $\tilde{x}$ is "nearby" data to original data x.

Note: A stable algorithm gives nearly the right answer to nearly the right question!

$\boxed{\text{Backward Stability}}$: (stronger type of stability). An algorithm $\tilde{f}$ for a problem f is backward stable if for each $x \in X$, $\tilde{f}(x) = f(\tilde{x})$ with $\dfrac{\|\tilde{x} - x\|}{\|x\|} = O(\epsilon_{mach})$.

Note: A Backward stable algorithm solves exactly a nearby problem.

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

answer from algorithm using data x

answer to original problem if original data x is changed to $\tilde{x} = x + \Delta x$. (nearby problem).

Note: This is the best we can hope for any algorithm.

Lesson 15

Page 111, Thm 15.1 : Suppose a backward stable algorithm is applied to solve a problem f with condition number κ on a computer that satisfies (13.5) and (13.7). Then, the relative errors satisfy

$$\frac{\| \tilde{f}(x) - f(x) \|}{\| f(x) \|} = O(\kappa(x) \cdot \epsilon_{mach})$$

Pf: (Read page 111)

15-2?

Pf: (of Thm 15.1)

Recall

$$K = \lim_{\delta \to 0} \sup_{\|\delta x\| \le \delta} \frac{\frac{\|\delta f\|}{\|f(x)\|}}{\frac{\|\delta x\|}{\|x\|}}$$

relative Change in Answer

relative change in data.

Where here $\delta f = f(\tilde{x}) - f(x)$ $\quad \tilde{x} = x + \delta x$

problem using data $\tilde{x}$

problem using data x.

This statement says

$$\frac{\|f(\tilde{x}) - f(x)\|}{\|f(x)\|} \le \left(K + o(1)\right) \frac{\|\tilde{x} - x\|}{\|x\|}$$

any term that $\to 0$ as $\delta \to 0$

$\|\tilde{x} - x\| \le \delta$

But for backward stable methods, $f(\tilde{x}) = \tilde{f}(x)$

$\Rightarrow$

$$\frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} \le \left(K + \underset{\text{little } o}{o(1)}\right) \frac{\|\tilde{x} - x\|}{\|x\|} \Rightarrow$$

$$\frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} = O\left(K \cdot \frac{\|\tilde{x} - x\|}{\|x\|}\right) = O(K \cdot \epsilon_{mach})$$

Since $\frac{\|\tilde{x} - x\|}{\|x\|} = O(\epsilon_{mach})$ for backward stable algs.