5374 Fall '22

Numerical Linear Algebra Lecture 7

Today:

* Algorithms in scientific computing
* Forward error, backward error, and condition numbers.
* Some more linear algebra review

References: Trefethen and Bau, lecture 12
Solomon, Chapter 2 section 2.2.

Note: Bring laptop on Thursday

More on the condition number of a matrix

Last time we introduced the condition number of a $n \times n$ matrix A

$$\operatorname{cond}(A) = \begin{cases} \|A\| \, \|A^{-1}\| & \text{if } A \text{ is invertible} \\ +\infty & \text{otherwise} \end{cases}$$

If $\overset{(n \times n)}{A}$ is invertible, $b \in \mathbb{R}^n$ then let z be the unique solution of

$$Az = b$$

We saw last time that if we choose a matrix $\dot{A}$ and a vector $\dot{b}$ and define (for ε small)

$$A_\varepsilon = A + \varepsilon \dot{A}$$
$$b_\varepsilon = b + \varepsilon \dot{b}$$
$$z_\varepsilon = A_\varepsilon^{-1} b_\varepsilon$$

The norm for matrices here is the operator norm $\|A\| = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$

Then

$$\frac{\|z - z_\varepsilon\|_2}{\|z\|_2} \leq \operatorname{cond}(A)\left(\frac{\|A - A_\varepsilon\|}{\|A\|} + \frac{\|b - b_\varepsilon\|_2}{\|b\|_2}\right) + O(\varepsilon^2)$$

Today, as a kind of warmup, let us show how this inequality can be improved if $\dot{A} = 0$; so $A_\varepsilon = A$ for all ε.

In this case we have z and z_ε which solve, respectively

$$Az = b, \quad Az_\varepsilon = b_\varepsilon$$

Without knowing z, and knowing z_ε, I would like to estimate how far away is

z_ε from z.

We cannot compute $\|z - z_\varepsilon\|_2$, but we can compute

$$\|Az - Az_\varepsilon\|_2$$
$$= \|b - Az_\varepsilon\|_2$$

Conversely, $z = A^{-1}b$, $z_\varepsilon = A^{-1}b_\varepsilon$, we have

$$\|z - z_\varepsilon\|_2 = \|A^{-1}(b - b_\varepsilon)\|_2$$

$$\leq \|A^{-1}\| \, \|b - b_\varepsilon\|_2$$

Moreover,

$$\frac{\|z - z_\varepsilon\|_2}{\|z\|_2} \leq \frac{\|A^{-1}\| \, \|b - b_\varepsilon\|_2}{\|z\|_2}$$

Since $b = Az$, $\|b\|_2 \leq \|A\| \, \|z\|_2$, implying that $\dfrac{1}{\|z\|_2} \leq \dfrac{\|A\|}{\|b\|_2}$

so,

$$\frac{\|z - z_\varepsilon\|_2}{\|z\|_2} \leq \|A\| \, \|A^{-1}\| \, \frac{\|b - b_\varepsilon\|_2}{\|b\|_2}$$

i.e.

$$\boxed{\frac{\|z - z_\varepsilon\|_2}{\|z\|_2}} \leq \operatorname{cond}(A) \boxed{\frac{\|b - b_\varepsilon\|_2}{\|b\|_2}}$$

This is called the (relative) forward error

This is called the (relative) backward error

This is an error we want to estimate but cannot compute directly (since we don't know z)

This quantity we can compute from knowledge of the problem and of our computed solution z_ε.

A framework for mathematical problems and algorithms to solve them

We will think of a mathematical problem simply as a function

f

$$X \xrightarrow{f} Y$$

(a subset of a vector space) (another subset of some vector space)

Every $x \in X$ will represent a particular instance, or a particular combination of parameters for the problem, and $f(x)$ will represent the solution of that problem for the instance x.

Examples: ① Consider the problem of computing the positive square root of a real positive number.

$$X = (0, \infty)$$
$$Y = (0, \infty)$$

and $f(x) = +\sqrt{x}$

② Consider the problem of, given A invertible $(n \times n)$ and $b \in \mathbb{R}^n$, to find z solving

$$Az = b$$

$$X = \{ (A,b) \mid A \text{ is an } n \times n \text{ invertible matrix}, \; b \in \mathbb{R}^n \}$$

$$\subset \mathbb{R}^{n^2+n}$$

$$Y = \mathbb{R}^n$$

$$f(A,b) = A^{-1} b$$

(3) Least square problems: Given A, a $m \times n$ matrix, and b, a vector in $\mathbb{R}^m$, find $z \in \mathbb{R}^n$ minimizing the norm

$$\|Az - b\|_2$$

$$X = \mathbb{R}^{mn+m}, \quad Y = \mathbb{R}^n$$

(4) Eigenvalue problems: Given a symmetric $n \times n$ matrix A, compute it's n eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_n$ sort in a non-decreasing order

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$$

$$X = \{ A \mid n \times n \text{ matrices with } A^t = A \}$$

$$Y = \{ y \in \mathbb{R}^n \mid y_1 \leq y_2 \leq \cdots \leq y_n \}$$

An algorithm is a more imprecise notion, it is a lot like a mathematical problem

$$f : X \to Y$$

except the function f comes with a set of instructions that take x as input and produces $f(x)$ in a finite number of steps.

To distinguish them from mathematical problems, we will denote algorithms with a $\wedge$ symbol, i.e. $\hat{f} : X \to Y$.

Our overarching goal is, given a mathematical problem $f : X \to Y$, to design or arrive at an algorithm $\hat{f} : X \to Y$ that serves as a good approximation for f.

How do we quantify how "good" an

algorithm is?

Backward error, forward error, and backward stability

(In the following discussion I will use $\| \cdot \|$ to refer to some unspecified norm in the space X, or in the space Y)

Fix a mathematical problem f, and an algorithm $\hat{f}$.

1. Forward error

$$\| f(x) - \hat{f}(x) \|$$

This is the error we most care about, but one we cannot compute directly since we lack knowledge of $f(x)$

(There is also the Relative forward error

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

2. Backward error

Here, given x, we look among all $\hat{x}$ in X, and try to find one such that

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

i.e. $\hat{f}(x)$ may not be the exact solution to problem x, but it is the exact solution of a (hopefully) close problem, $\hat{x}$.

That is, consider

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

(↑ this might be empty)

Suppose for a second that $f : X \to Y$ is Lipschitz continuous with norm L, i.e.

$$\|f(x) - f(y)\| \leq L \|x - y\| \quad \forall\, x, y \in X$$

If so, then

$$\underbrace{\|f(x) - \hat{f}(x)\|}_{\uparrow} \overset{\substack{\text{hopefully} \\ \downarrow}}{=} \|f(x) - f(\hat{x})\|$$

$$\leq L \|x - \hat{x}\|.$$

forward ξ error

$\underbrace{\qquad\qquad}_{\text{bounded by the backward error}}$