Computational Linear Algebra
Conditioning: A Measure of Error Magnification

David Shuman

February 10, 2022

Conditioning and Stability

- *Conditioning* refers to the perturbation behavior of a problem
  - Problem is *well-conditioned* if a small perturbation in the input data leads to a small perturbation in the output of the problem
  - Problem is *ill-conditioned* if a small perturbation in the input data leads to a large perturbation in the output of the problem

Conditioning and Stability

- *Conditioning* refers to the perturbation behavior of a problem
  - Problem is *well-conditioned* if a small perturbation in the input data leads to a small perturbation in the output of the problem
  - Problem is *ill-conditioned* if a small perturbation in the input data leads to a large perturbation in the output of the problem
- *Stability* refers to the perturbation behavior of a numerical algorithm to solve that problem on a computer

Forward and Backward Error (for Root-Finding Problem)

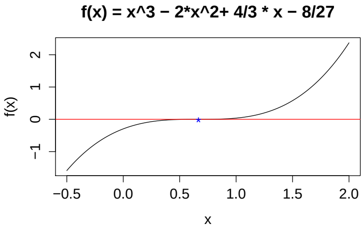
- Our problem solving process is like this. If r is the exact solution to $f(x) = 0$ and r_a is our approximate solution, then

$$\underbrace{\left\{\begin{array}{c}\text{data that defines}\\ \text{the problem:}\\ f(x)\end{array}\right\}}_{backward\ \ error:\ |f(r_a)|} \longrightarrow \left\{\begin{array}{c}\text{equation}\\ \text{solver}\end{array}\right\} \longrightarrow \underbrace{\left\{\begin{array}{c}r \text{ such that}\\ f(r) = 0\end{array}\right\}}_{forward\ \ error:\ |r - r_a|}$$

- The forward error is the difference between the exact answer and the computed answer $|r - r_a|$
- The computed answer r_a is the solution to some perturbed problem $\bar{f}(x) = 0$. Thus $\bar{f}(r_a) = 0$ with $\bar{f} \approx f$
- The backward error is the difference between the original problem and the perturbed problem: $|f(r_a) - \bar{f}(r_a)| = |f(r_a)|$. In some sense tells us how far we are from the problem that the algorithm actually solved

Forward and Backward Error (for Root-Finding Problem)

- If r_a is our approximate solution to $f(x) = 0$ and r is the exact solution, then
  forward error: $|r - r_a|$ backward error: $|f(r_a)|$
- In Example 1.7



$$r = 2/3, \quad r_a = 0.6666641, \quad \text{and} \quad f(r_a) = 0$$

- forward error: $|r - r_a| = 2.543132 * 10^{-6}$
- backward error: $|f(r_a)| = 0$ (less than ε_M)

Condition Number (informally)

- The condition number is the maximum error magnification:
  (forward error) $\leq$ (condition number) $*$ (backward error).
- If the condition number is big, then small changes in the problem (backward error) lead to big errors in the solution (forward error)
- That is, how much small changes in the problem are magnified
- A problem with large condition number is ill-conditioned
- The condition number of root finding is proportional to $\frac{1}{f'(r)}$
- Note: Conditioning can be different at different points!

Reminder: Matrix Norms

Def. The Matrix p-norm is given by

$$||A||_p = \max_{\vec{x} \neq \vec{0}} \frac{||A\vec{x}||_p}{||\vec{x}||_p} = \max_{||\vec{x}||_p = 1} ||A\vec{x}||_p$$

- $||A||_p$ gives the maximum relative expansion by A
- Apply A to the unit sphere in $\mathbb{R}^n$. Then $||A||_p$ is the length of the vector that is farthest from the origin in the image
- Key property: $||A\vec{x}||_p \leq ||A||_p ||\vec{x}||_p$

Error Magnification in Solving $A\vec{x} = \mathrm{b}$

Suppose we solve $A\vec{x} = \mathrm{b}$ and get an approximate answer $\vec{x}_a$

	Backward Error	Forward Error
Absolute	$\|\|\mathrm{b} - A\vec{x}_a\|\|$	$\|\|\vec{x} - \vec{x}_a\|\|$
Relative	$\frac{\|\|\mathrm{b}-A\vec{x}_a\|\|}{\|\|\mathrm{b}\|\|}$	$\frac{\|\|\vec{x}-\vec{x}_a\|\|}{\|\|\vec{x}\|\|}$

The error magnification is the number given by

$$\text{(relative backward error)} \times \text{(error magnification)} = \text{(relative forward error)}$$

or

$$\text{(error magnification)} = \frac{\text{(relative forward error)}}{\text{(relative backward error)}} = \frac{\left(\frac{||\vec{x}-\vec{x}_a||}{||\vec{x}||}\right)}{\left(\frac{||\mathrm{b}-A\vec{x}_a||}{||\mathrm{b}||}\right)}$$

Book's Example 2.11 on board

The Condition Number of a Matrix

Def. The condition number $\text{Cond}(A) = \kappa(A)$ is the lowest upper bound on the error magnification for solving $A\vec{x} = \text{b}$

Thm. For an $n \times n$ matrix A, $\quad \text{Cond}(A) = ||A|| \cdot ||A^{-1}||$

- Note how it depends on the matrix norm, and it bounds the error magnification in the corresponding vector norm
- Computing the inverse is hard, so many algorithms give a "cheap" estimate of the condition number without computing it exactly
- In today's activity, I've provided a function Cond that computes it
- Read the help menu on R's built in function: `kappa`

Eg. Compute the condition number for various p-norms for

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 \\ 1.0001 & 1 \end{pmatrix}, \qquad I_n,$$

Diagonal Matrices,
Hilbert Matrices H_n

Properties of the Condition Number

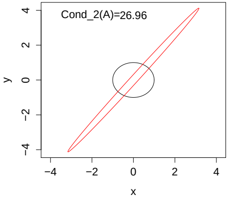
- In solving $A\vec{x} = \text{b}$: $\quad$ (Rel FE) $\leq \text{Cond}(A) \cdot$ (Rel BE)
- $\text{Cond}(A) \geq 1$ for any matrix A. $\text{Cond}(\alpha I) = 1$
- $\text{Cond}(AB) \leq \text{Cond}(A)\text{Cond}(B)$
- The problem $A\vec{x} = \text{b}$ is ill-conditioned if $\text{Cond}(A)$ is large (i.e., $10^5, 10^8, 10^{10}$, etc). Otherwise it is well-conditioned
- The condition number depends on the underlying norm
- $\text{Cond}(A) = ||A||_p \cdot ||A^{-1}||_p = \dfrac{\left(\max\limits_{||\vec{x}||_p=1} ||A\vec{x}||\right)}{\left(\min\limits_{||\vec{x}||_p=1} ||A\vec{x}||_p\right)}$

  This measures
  - the maximum relative expansion by A divided by the minimum relative expansion by A
  - the eccentricity of the image of the unit sphere by A
- If $\text{Cond}(A)$ is large, then A is close to singular

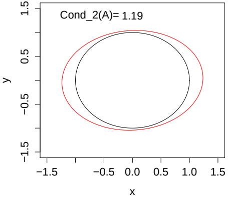
Examples of the Condition Number for 2x2 Matrices

- Image of the unit sphere in the 2-norm under a linear mapping A is a hyperellipse
- Using the 2-norm, the condition number of A, $\kappa_2(A)$ gives the ratio of the length of the longest principal semiaxis to the length of the shortest principal semiaxis (maximum and minimum singular values)

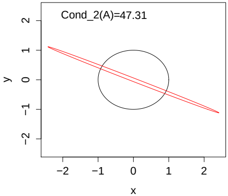
$$\begin{pmatrix} 3 & 1 \\ 4 & 1 \end{pmatrix}$$



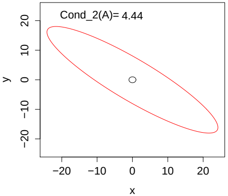
$$\begin{pmatrix} 0.3 & 1.2 \\ -1.0 & 0.3 \end{pmatrix}$$



$$\begin{pmatrix} 2.1 & -1.2 \\ -1.0 & 0.5 \end{pmatrix}$$



$$\begin{pmatrix} 21 & -12 \\ -10 & 15 \end{pmatrix}$$



Appendix: Another Numerical Example

- Consider $\begin{pmatrix} 0.913 & 0.659 \\ 0.457 & 0.330 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.254 \\ 0.127 \end{pmatrix}$
- The condition number is: $\mathrm{Cond}_2(A) = 1.25 \times 10^4$
- The exact solution is $\vec{x} = (1, -1)$.
- Consider the two approximate solutions (2-norm)

$$
\begin{aligned}
\vec{x}_1 &= (-0.0827, 0.5) \\
\triangle\vec{x}_1 &= (1.0827, -1.5) \\
||\triangle\vec{x}_1|| &= 1.85 \\
||\triangle\vec{x}_1||/||\vec{x}|| &= 1.308 \\
b_1 &= (0.2539949, 0.1272061) \\
\triangle b_1 &= (0.0000051, -0.0002061) \\
||\triangle b_1|| &= 0.000206 \\
||\triangle b_1||/||b|| &= 0.000726 \\
mag &= 1.8 \times 10^3
\end{aligned}
$$

$$
\begin{aligned}
\vec{x}_2 &= (0.999, -1.001) \\
\triangle\vec{x}_2 &= (0.001, 0.001) \\
||\triangle\vec{x}_2|| &= .0014 \\
||\triangle\vec{x}_2||/||\vec{x}|| &= .001 \\
b_2 &= (0.252428, 0.126213) \\
\triangle b_2 &= (0.001572, 0.000787) \\
||\triangle b_2|| &= .00176 \\
||\triangle b_2||/||b|| &= .0062 \\
mag &= 1.6 \times 10^1
\end{aligned}
$$

- the condition number bounds the magnification
- with large condition numbers, a small residual does not imply a small error in the solution

Appendix: Conditioning of the Root-Finding Problem

- Problem: $f(x) = 0$ with root r
- Small change $\epsilon g(x)$ is made to input (a function, not a variable)
- Perturbed problem: $f(x) + \epsilon g(x) = 0$
- Root of perturbed problem is $r + \Delta r$: $f(r + \Delta r) + \epsilon g(r + \Delta r) = 0$
- Expand f and g in degree-one Taylor polynomials:

$$f(r) + (\Delta r)f'(r) + \epsilon g(r) + \epsilon(\Delta r)g'(r) + O((\Delta r)^2) = 0$$

- For small Δr, ignore higher-order term:

$$(\Delta r)(f'(r) + \epsilon g'(r)) \approx -f(r) - \epsilon g(r) = -\epsilon g(r)$$

- So for ϵ small and $f'(r) \neq 0$:

$$(\Delta r) \approx \frac{-\epsilon g(r)}{f'(r) + \epsilon g'(r)} \approx \frac{-\epsilon g(r)}{f'(r)}$$

- $|\text{Forward error}/\text{backward error}| = |\Delta r|/|\epsilon g(r)| \approx 1/|f'(r)|$
- Can be used to analyze how errors propagate from the input to the output (see activity A3, Ex. 2, part c and Example 1.10 from the textbook)