Errors in Linear Systems Solving

CS/SE 4X03

Ned Nedialkov

McMaster University

September 30, 2021

Outline

Norms

Residual

Relative solution error

Norms

Vector norms

Norm is a function $\| \cdot \|$ that satisfies for any $x \in \mathbb{R}^n$

1. $\|x\| \geq 0$, and $\|x\| = 0$ iff $x = 0$, the zero vector
2. $\|\alpha x\| = |\alpha| \|x\|$, $\alpha \in \mathbb{R}$
3. $\|x + y\| \leq \|x\| + \|y\|$ for any $x, y \in \mathbb{R}^n$

lp norms

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}, \qquad 1 \leq p \leq \infty$$

Norms cont.

- $p = 1$, one norm

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

- $p = \infty$, infinity or max norm

$$\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$$

- $p = 2$, two or Euclidean norm

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Norms cont.

Matrix norms

- $A \in \mathbb{R}^{m \times n}$, $\| \cdot \|$ is a vector norm
- Matrix norm induced by this vector norm

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

- Properties
  1. $\|A\| \geq 0$, and $\|A\| = 0$ iff $A = 0$, the zero matrix
  2. $\|\alpha A\| = |\alpha| \|A\|$, $\alpha \in \mathbb{R}$
  3. $\|A + B\| = \|A\| + \|B\|$, for any $A, B \in \mathbb{R}^{m \times n}$
  4. $\|AB\| \leq \|A\| \cdot \|B\|$, for any $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$

- Infinity norm, max row sum

$$\|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$$

- One norm, max column sum

$$\|A\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$$

- Two norm

$$\|A\|_2 = \max_i \sqrt{\lambda_i(A^T A)},$$

where $\lambda_i(A^T A)$ is the ith eigenvalue of $A^T A$

Residual

Consider $Ax = b$

- Let $\widetilde{x}$ be the computed solution, and let x be the exact solution
- Relative error in the solution is

$$\frac{\|x - \widetilde{x}\|}{\|x\|}$$

- Residual is

$$r = b - A\widetilde{x}$$

$$r = 0 \iff b - A\widetilde{x} = 0 \iff \widetilde{x} = x$$

- In practice $r \neq 0$

- $Ax = b$ and $\alpha Ax = \alpha b$ have the same solution
  α is a scalar
- $r_\alpha = \alpha b - \alpha A\widetilde{x} = \alpha(b - A\widetilde{x})$ can be arbitrarily large
- residual can be arbitrarily large

Residual cont.

Example 1. Consider

$$A = \begin{bmatrix} 1.2969 & 0.8648 \\ 0.2161 & 0.1441 \end{bmatrix}, \qquad b = \begin{bmatrix} 0.8642 \\ 0.1440 \end{bmatrix}$$

and the approximate solution $\widetilde{x} = [0.9911, -0.487]^T$

- The residual is small:

$$r = b - A\widetilde{x} \approx [10^{-8}, -10^{-8}]^T, \qquad \|r\|_\infty \approx 10^{-8}$$

- The exact solution is $x = [2,\ -2]^T$. The error in $\widetilde{x}$ is large:

$$x - \widetilde{x} = [1.513, -1.0089], \qquad \|x - \widetilde{x}\|_\infty = 1.513$$

- Small residual does not imply small solution error

Relative solution error

Given $\widetilde{x}$, how large is

$$\frac{\|x - \widetilde{x}\|}{\|x\|} \tag{1}$$

Using $r = b - A\widetilde{x} = Ax - A\widetilde{x} = A(x - \widetilde{x})$,

$$\begin{aligned} x - \widetilde{x} &= A^{-1}r \\ \|x - \widetilde{x}\| &= \|A^{-1}r\| \leq \|A^{-1}\|\|r\| \end{aligned} \tag{2}$$

Using $b = Ax$, $\|b\| = \|Ax\| \leq \|A\|\|x\|$, and

$$\|x\| \geq \frac{\|b\|}{\|A\|} \tag{3}$$

The condition number of A is

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

Using (2–3) in (1),

$$\frac{\|x - \widetilde{x}\|}{\|x\|} \leq \frac{\|A^{-1}\| \|r\|}{\frac{\|b\|}{\|A\|}} = \|A^{-1}\| \|A\| \frac{\|r\|}{\|b\|} = \text{cond}(A) \frac{\|r\|}{\|b\|}$$

$$\frac{\|x - \widetilde{x}\|}{\|x\|} \leq \text{cond}(A) \frac{\|r\|}{\|b\|}$$

- If $\text{cond}(A)$ is not large and $\|r\|/\|b\|$ is small then small relative error
- As a rule of thumb, if $\text{cond}(A) \approx 10^k$, then about k decimal digits are lost when solving $Ax = b$.

- In our example

$$A^{-1} = 10^8 \begin{bmatrix} 0.1441 & -0.8648 \\ -0.2161 & 1.2869 \end{bmatrix}$$

- In the two norm, $\text{cond}(A) \approx 2.4973 \cdot 10^8$

$$\text{cond}(A)\frac{\|r\|}{\|b\|} \approx 4.0311$$

$$\frac{\|x - \widetilde{x}\|}{\|x\|} \approx 0.6429$$