You Perturb It, You Pay For It According to $\mathcal{K}(A)$: Norms and Error Analysis

Nate DeMaagd, Kurt O'Hearn

MTH 499-02

March 25, 2013

Outline

- Norms: properties, types, examples

Outline

- Norms: properties, types, examples
- Theorems, results on norms

- Norms: properties, types, examples
- Theorems, results on norms
- Relative/absolute error in perturbations, conditioning number

Definition

Let V be a vector space. A norm is a function $\|\cdot\| : V \to \mathbb{R}^+$ that has the following properties:

- $\|x\| > 0$ if $x \neq 0, \ x \in V$
- $\|\lambda x\| = |\lambda| \ \|x\|$ if $\lambda \in \mathbb{R}, \ x \in V$
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ (triangle inequality)

Definition

Let V be a vector space. A norm is a function $\|\cdot\| : V \to \mathbb{R}^+$ that has the following properties:

- $\|x\| > 0$ if $x \neq 0,\ x \in V$
- $\|\lambda x\| = |\lambda|\ \|x\|$ if $\lambda \in \mathbb{R},\ x \in V$
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ (triangle inequality)

Other comments:

- Norm can be thought of as generalization of absolute value
- Interpretation: vector length in $\mathbb{R}^2$, $\mathbb{R}^3$

Vector Norms:

Vector Norms:

- l_k norm

$$l_k = \|x\|_k = \left(\sum_{i=1}^{n} |x_i|^k \right)^{1/k} \text{ where } k \in \{1, 2, \dots\} \text{ and } x = (x_1, \dots, x_n)^T$$

Vector Norms:

- l_k norm

$$l_k = \|x\|_k = \left(\sum_{i=1}^{n} |x_i|^k \right)^{1/k} \quad \text{where } k \in \{1, 2, \dots\} \text{ and } x = (x_1, \dots, x_n)^T$$

- Some common instances of the l_k norm:

k	l_k
1	$\|x\|_1 = \sum_{i=1}^{n} \lvert x_i \rvert$
2	$\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$
∞	$\|x\|_\infty = \max_{1 \le i \le n} \lvert x_i \rvert$

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\| \cdot \|\|_1$	$\|\| \cdot \|\|_2$	$\|\| \cdot \|\|_\infty$
x			
y			
z			

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\|\cdot\|\|_1$	$\|\|\cdot\|\|_2$	$\|\|\cdot\|\|_\infty$
x	6		
y			
z			

$$||x||_1 = 2 + 1 + 3 = 6$$

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\cdot\|_1$	$\|\cdot\|_2$	$\|\cdot\|_\infty$
x	6	3.74	
y			
z			

$$||x||_2 = \sqrt{2^2 + 1^2 + 3^2} = 3.74$$

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\|\cdot\|\|_1$	$\|\|\cdot\|\|_2$	$\|\|\cdot\|\|_\infty$
x	6	3.74	3
y			
z			

$$||x||_\infty = 3$$

- Consider vectors in $\mathbb{R}^3$

$$x = (2, -1, 3)^T \qquad y = (7, 2, 5)^T \qquad z = (-3, 1, -9)^T$$

	$\|\cdot\|_1$	$\|\cdot\|_2$	$\|\cdot\|_\infty$
x	6	3.74	3
y	14	8.83	7
z	13	9.54	9

Definition

A matrix norm subordinate to a vector norm $\|\cdot\|$ is defined as

$$\|A\| = \sup\{\|Au\| : u \in \mathbb{R}^n, \|u\| = 1\},$$

where A is an $n \times n$ matrix.

Definition

A matrix norm subordinate to a vector norm $\|\cdot\|$ is defined as

$$\|A\| = \sup\left\{\|Au\| : u \in \mathbb{R}^n, \|u\| = 1\right\},$$

where A is an $n \times n$ matrix.

- Two common matrix norms

Definition

A matrix norm subordinate to a vector norm $\|\cdot\|$ is defined as

$$\|A\| = \sup\left\{\|Au\| : u \in \mathbb{R}^n, \|u\| = 1\right\},$$

where A is an $n \times n$ matrix.

- Two common matrix norms
  - Subordinate to vector norm $\|\cdot\|_\infty$ is

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|,$$

which is the maximum absolute row sum of A.

Definition

A matrix norm subordinate to a vector norm $\|\cdot\|$ is defined as

$$\|A\| = \sup \{\|Au\| : u \in \mathbb{R}^n, \|u\| = 1\},$$

where A is an $n \times n$ matrix.

- Two common matrix norms
  - Subordinate to vector norm $\|\cdot\|_\infty$ is

    $$\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|,$$

    which is the maximum absolute row sum of A.
  - Subordinate to vector norm $\|\cdot\|_2$ is

    $$\begin{aligned} \|A\|_2 &= \sup_{\|x\|_2 = 1} \{\|Ax\|_2\} \\ &= \sqrt{\rho(A^T A)}, \end{aligned}$$

    where $\rho(A^T A)$ is spectral radius of $A^T A$, largest eigenvalue of $A^T A$.

- Consider matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$

- Consider matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$
- $(A^T A) = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$

- Consider matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$
- $(A^T A) = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$
- For $A^T A$, eigenvalues are $\sigma_1^2 = 6.8541$ and $\sigma_2^2 = 0.1459$.

- Consider matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$
- $(A^T A) = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$
- For $A^T A$, eigenvalues are $\sigma_1^2 = 6.8541$ and $\sigma_2^2 = 0.1459$.
- So, $\|A\|_2 = \sqrt{6.8541} \approx 2.618$.

- Consider the matrix $B = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

- Consider the matrix $B = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$
- $\|B\|_2$ is simply the maximum absolute row sum, which is 9.

Theorem (Subordinate Matrix Norm)

If $\|\cdot\|$ is any norm in $\mathbb{R}^n$, then the equation

$$\|A\| = \sup_{\|u\|=1} \{\|Au\| : u \in \mathbb{R}^n\}$$

defines a norm on the linear space of all $n \times n$ matrices.

Theorem (Subordinate Matrix Norm)

If $\|\cdot\|$ is any norm in $\mathbb{R}^n$, then the equation

$$\|A\| = \sup_{\|u\|=1} \{\|Au\| : u \in \mathbb{R}^n\}$$

defines a norm on the linear space of all $n \times n$ matrices.

Proof.

Follows from properties of vector norms. Questions? $\square$

Theorem (Subordinate Matrix Norm)

If $\|\cdot\|$ is any norm in $\mathbb{R}^n$, then the equation

$$\|A\| = \sup_{\|u\|=1} \{\|Au\| : u \in \mathbb{R}^n\}$$

defines a norm on the linear space of all $n \times n$ matrices.

Proof.

Follows from properties of vector norms. Questions? $\square$

Other Results

- $\|Ax\| \leq \|A\|\|x\| \qquad (x \in \mathbb{R}^n)$
- $\|AB\| \leq \|A\|\|B\|$
- $\|I\| = 1$

Results

- Let $Ax = b$ be a linear system and let B be a perturbation of A such that $B\tilde{x} = b$. Then the following holds:

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \|I - BA\|$$

Results

- Let $Ax = b$ be a linear system and let B be a perturbation of A such that $B\tilde{x} = b$. Then the following holds:

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \|I - BA\|$$

- Let $Ax = b$ be a linear system and let $\tilde{b}$ be a perturbation of b such that $A\tilde{x} = \tilde{b}$. Then the following holds:

$$\|x - \tilde{x}\| \leq \|A^{-1}\| \, \|b - \tilde{b}\|$$

Results

- Let $Ax = b$ be a linear system and let B be a perturbation of A such that $B\tilde{x} = b$. Then the following holds:

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \|I - BA\|$$

- Let $Ax = b$ be a linear system and let $\tilde{b}$ be a perturbation of b such that $A\tilde{x} = \tilde{b}$. Then the following holds:

$$\|x - \tilde{x}\| \leq \|A^{-1}\| \, \|b - \tilde{b}\|$$

- Putting these together, we get

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \mathcal{K}(A) \frac{\|b - \tilde{b}\|}{\|b\|}$$

where $\mathcal{K}(A) = \|A\| \, \|A^{-1}\|$ is called the condition number of A

Results

- Let $Ax = b$ be a linear system and let B be a perturbation of A such that $B\tilde{x} = b$. Then the following holds:

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \|I - BA\|$$

- Let $Ax = b$ be a linear system and let $\tilde{b}$ be a perturbation of b such that $A\tilde{x} = \tilde{b}$. Then the following holds:

$$\|x - \tilde{x}\| \leq \|A^{-1}\| \, \|b - \tilde{b}\|$$

- Putting these together, we get

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \mathcal{K}(A) \frac{\|b - \tilde{b}\|}{\|b\|}$$

  where $\mathcal{K}(A) = \|A\| \, \|A^{-1}\|$ is called the condition number of A

- $\mathcal{K}(A)$ large, A ill-conditioned, else A well-conditioned

Results

- Let $Ax = b$ be a linear system and let B be a perturbation of A such that $B\tilde{x} = b$. Then the following holds:

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \|I - BA\|$$

- Let $Ax = b$ be a linear system and let $\tilde{b}$ be a perturbation of b such that $A\tilde{x} = \tilde{b}$. Then the following holds:

$$\|x - \tilde{x}\| \leq \|A^{-1}\| \, \|b - \tilde{b}\|$$

- Putting these together, we get

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \mathcal{K}(A) \frac{\|b - \tilde{b}\|}{\|b\|}$$

  where $\mathcal{K}(A) = \|A\| \, \|A^{-1}\|$ is called the condition number of A
- $\mathcal{K}(A)$ large, A ill-conditioned, else A well-conditioned
- $\mathcal{K}(A) > 0$

Definition

We define the error e and residual r for perturbed systems as follows:

$$e = x - \tilde{x} \quad \text{and} \quad r = b - \tilde{b} = Ax - A\tilde{x} = Ae$$

Definition

We define the error e and residual r for perturbed systems as follows:

$$e = x - \tilde{x} \quad \text{and} \quad r = b - \tilde{b} = Ax - A\tilde{x} = Ae$$

Results

- For $Ax = b$, we have

$$\frac{1}{\mathcal{K}(A)} \frac{||r||}{||b||} \leq \frac{||e||}{||x||} \leq \mathcal{K}(A) \frac{||r||}{||b||}$$

Definition

We define the error e and residual r for perturbed systems as follows:

$$e = x - \tilde{x} \quad \text{and} \quad r = b - \tilde{b} = Ax - A\tilde{x} = Ae$$

Results

- For $Ax = b$, we have

$$\frac{1}{\mathcal{K}(A)} \frac{\|r\|}{\|b\|} \leq \frac{\|e\|}{\|x\|} \leq \mathcal{K}(A) \frac{\|r\|}{\|b\|}$$

- Question: why useful?

Definition

We define the error e and residual r for perturbed systems as follows:

$$e = x - \tilde{x} \quad \text{and} \quad r = b - \tilde{b} = Ax - A\tilde{x} = Ae$$

Results

- For $Ax = b$, we have

$$\frac{1}{\mathcal{K}(A)} \frac{||r||}{||b||} \leq \frac{||e||}{||x||} \leq \mathcal{K}(A) \frac{||r||}{||b||}$$

- Question: why useful?
- Answer: b, r, $\mathcal{K}(A)$ known, so can estimate the relative error in x

Consider