AMS526: Numerical Analysis I (Numerical Linear Algebra)

Lecture 4: Matrix Norms

Xiangmin Jiao

SUNY Stony Brook

September 11, 2008

Review

- Orthogonal projection matrix
- Let $\boldsymbol{Q}$ be composed of column vectors $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n\}$.
    - Then $\boldsymbol{Q}\boldsymbol{Q}^* = \sum_{i=1}^{n}(\boldsymbol{q}_i\boldsymbol{q}_i^*)$ is an orthogonal projection matrix.
    - In addition, $\boldsymbol{I} - \boldsymbol{Q}\boldsymbol{Q}^*$ is also an orthogonal projection matrix.
- Unitary matrix: $\boldsymbol{U}\boldsymbol{U}^* = \boldsymbol{U}^*\boldsymbol{U} = \boldsymbol{I}$.
- Vector norm, p-norm, weighted p-norm

Outline

General Matrix Norms

- One can view $m \times n$ matrices as mn-dimensional vectors and obtain *general matrix norms*, which satisfy (for $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{m \times n}$)

$$
\begin{aligned}
&(1)\|\boldsymbol{A}\| \geq 0, \text{ and } \|\boldsymbol{A}\| = 0 \text{ only if } \boldsymbol{A} = \boldsymbol{0}, \\
&(2)\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|, \\
&(3)\|\alpha \boldsymbol{A}\| = |\alpha| \|\boldsymbol{A}\|.
\end{aligned}
$$

General Matrix Norms

- One can view $m \times n$ matrices as mn-dimensional vectors and obtain *general matrix norms*, which satisfy (for $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{m\times n}$)

$$(1)\|\boldsymbol{A}\| \geq 0, \text{ and } \|\boldsymbol{A}\| = 0 \text{ only if } \boldsymbol{A} = \boldsymbol{0},$$
$$(2)\|\boldsymbol{A}+\boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|,$$
$$(3)\|\alpha\boldsymbol{A}\| = |\alpha|\|\boldsymbol{A}\|.$$

- One useful norm is Frobenius norm (a.k.a. Hilbert-Schmidt norm)

$$\|\boldsymbol{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} = \sqrt{\sum_{j=1}^{n} \|\boldsymbol{a}_j\|_2^2}$$

  i.e., 2-norm of mn-vector
- Furthermore,

$$\|\boldsymbol{A}\|_F = \sqrt{\text{tr}(\boldsymbol{A}^*\boldsymbol{A})}$$

  where $\text{tr}(\boldsymbol{B})$ denotes trace of $\boldsymbol{B}$, the sum of its diagonal entries

Matrix Norms Induced by Vector Norms

- Viewing $m \times n$ matrix as mn-vectors is not always useful, as operations involving $m \times n$ matrices do not behave this way
- *Induced matrix norms* capture such behavior

Matrix Norms Induced by Vector Norms

- Viewing $m \times n$ matrix as mn-vectors is not always useful, as operations involving $m \times n$ matrices do not behave this way
- *Induced matrix norms* capture such behavior

Definition

Given vector norms $\|\cdot\|_{(n)}$ and $\|\cdot\|_{(m)}$ on domain and range of $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, respectively, the induced matrix norm $\|A\|_{(m,n)}$ is the smallest number $C \in \mathbb{R}$ for which the following inequality holds for all $\boldsymbol{x} \in \mathbb{C}^n$:

$$\|\boldsymbol{A}\boldsymbol{x}\|_{(m)} \leq C\|\boldsymbol{x}\|_{(n)}.$$

Matrix Norms Induced by Vector Norms

- Viewing $m \times n$ matrix as mn-vectors is not always useful, as operations involving $m \times n$ matrices do not behave this way
- *Induced matrix norms* capture such behavior

Definition

Given vector norms $\|\cdot\|_{(n)}$ and $\|\cdot\|_{(m)}$ on domain and range of $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, respectively, the induced matrix norm $\|A\|_{(m,n)}$ is the smallest number $C \in \mathbb{R}$ for which the following inequality holds for all $\boldsymbol{x} \in \mathbb{C}^n$:

$$\|\boldsymbol{A}\boldsymbol{x}\|_{(m)} \leq C\|\boldsymbol{x}\|_{(n)}.$$

- In other words, it is supremum of ratio $\|\boldsymbol{A}\boldsymbol{x}\|_{(m)}/\|\boldsymbol{x}\|_{(n)}$ for all nonzero vectors $\boldsymbol{x} \in \mathbb{C}^n$
- Maximum factor by which $\boldsymbol{A}$ can "stretch" $\boldsymbol{x} \in \mathbb{C}^n$

$$\|\boldsymbol{A}\|_{(m,n)} = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \boldsymbol{x}\neq\boldsymbol{0}} \|\boldsymbol{A}\boldsymbol{x}\|_{(m)}/\|\boldsymbol{x}\|_{(n)} = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \|\boldsymbol{x}\|_{(n)}=1} \|\boldsymbol{A}\boldsymbol{x}\|_{(m)}.$$

Matrix Norms Induced by Vector Norms

- Viewing $m \times n$ matrix as mn-vectors is not always useful, as operations involving $m \times n$ matrices do not behave this way
- *Induced matrix norms* capture such behavior

Definition

Given vector norms $\|\cdot\|_{(n)}$ and $\|\cdot\|_{(m)}$ on domain and range of $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, respectively, the induced matrix norm $\|A\|_{(m,n)}$ is the smallest number $C \in \mathbb{R}$ for which the following inequality holds for all $\boldsymbol{x} \in \mathbb{C}^n$:

$$\|\boldsymbol{A}\boldsymbol{x}\|_{(m)} \leq C\|\boldsymbol{x}\|_{(n)}.$$

- In other words, it is supremum of ratio $\|\boldsymbol{A}\boldsymbol{x}\|_{(m)}/\|\boldsymbol{x}\|_{(n)}$ for all nonzero vectors $\boldsymbol{x} \in \mathbb{C}^n$
- Maximum factor by which $\boldsymbol{A}$ can "stretch" $\boldsymbol{x} \in \mathbb{C}^n$

$$\|\boldsymbol{A}\|_{(m,n)} = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \boldsymbol{x}\neq\boldsymbol{0}} \|\boldsymbol{A}\boldsymbol{x}\|_{(m)}/\|\boldsymbol{x}\|_{(n)} = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \|\boldsymbol{x}\|_{(n)}=1} \|\boldsymbol{A}\boldsymbol{x}\|_{(m)}.$$

- Is vector norm consistent with matrix norm of $m \times 1$-matrix?

1-norm

- By definition

$$\|\boldsymbol{A}\|_1 = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1$$

1-norm

- By definition

$$\|\boldsymbol{A}\|_1 = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1$$

- What is it equal to?

1-norm

- By definition

$$\|\boldsymbol{A}\|_1 = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1$$

- What is it equal to?
  - Maximum of 1-norm of column vectors of $\boldsymbol{A}$
  - "maximum column sum" of $\boldsymbol{A}$ is oversimplified in the textbook

1-norm

- By definition

$$\|\boldsymbol{A}\|_1 = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1$$

- What is it equal to?
  - Maximum of 1-norm of column vectors of $\boldsymbol{A}$
  - "maximum column sum" of $\boldsymbol{A}$ is oversimplified in the textbook
- To show it, note that for $\boldsymbol{x} \in \mathbb{C}^n$ and $\|\boldsymbol{x}\|_1 = 1$

$$\|\boldsymbol{A}\boldsymbol{x}\|_1 = \|\sum_{j=1}^{n} x_j \boldsymbol{a}_j\|_1 \leq \sum_{j=1}^{n} |x_j| \|\boldsymbol{a}_j\|_1 \leq \max_{1\leq j\leq n} \|\boldsymbol{a}_j\|_1 \|\boldsymbol{x}\|_1$$

- Let $k = \arg\max_{1\leq j\leq n} \|\boldsymbol{a}_j\|_1$, then $\|\boldsymbol{A}\boldsymbol{e}_k\|_1 = \|\boldsymbol{a}_k\|_1$, so $\max_{1\leq j\leq n} \|\boldsymbol{a}_j\|_1$ is tight upper bound

∞-norm

- By definition
$$\|\boldsymbol{A}\|_\infty = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \|\boldsymbol{x}\|_\infty=1} \|\boldsymbol{Ax}\|_\infty$$
- What is $\|\boldsymbol{A}\|_\infty$ equal to?

∞-norm

- By definition
$$\|\boldsymbol{A}\|_\infty = \sup_{\boldsymbol{x}\in\mathbb{C}^n, \|\boldsymbol{x}\|_\infty=1} \|\boldsymbol{Ax}\|_\infty$$
- What is $\|\boldsymbol{A}\|_\infty$ equal to?
  - Maximum of 1-norm of column vectors of $\boldsymbol{A}^*$
- To show it, note that for $\boldsymbol{x} \in \mathbb{C}^n$ and $\|\boldsymbol{x}\|_\infty = 1$
$$\|\boldsymbol{Ax}\|_\infty = \max_{1\leq i\leq m} |\boldsymbol{a}_i^* x| \leq \max_{1\leq i\leq m} \|\boldsymbol{a}_i^*\|_1 \|\boldsymbol{x}\|_\infty$$
  where $\boldsymbol{a}_i^*$ denotes ith row vector of $\boldsymbol{A}$
- Furthermore, $\max_{1\leq i\leq m} \|\boldsymbol{a}_i^*\|_1$ is a tight bound.
  - Which vector can we choose to reach the bound?

2-norm

- What is 2-norm of a matrix?

2-norm

- What is 2-norm of a matrix?
- Answer: Its largest singular value.
- We will talk more about singular-value decomposition

2-norm

- What is 2-norm of a matrix?
- Answer: Its largest singular value.
- We will talk more about singular-value decomposition
- What is 2-norm of a diagonal matrix?

Cauchy-Schwarz and Hölder Inequalities

- Hölder inequality: Let p and q satisfy $1/p + 1/q = 1$ with $1 \leq p, q \leq \infty$, then

$$|\boldsymbol{x}^* \boldsymbol{y}| \leq \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q$$

- Cauchy-Schwarz inequality

$$|\boldsymbol{x}^* \boldsymbol{y}| \leq \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2$$

- Cauchy-Schwarz inequality is a special case of Hölder inequality

Cauchy-Schwarz and Hölder Inequalities

- Hölder inequality: Let p and q satisfy $1/p + 1/q = 1$ with $1 \leq p, q \leq \infty$, then
$$|\boldsymbol{x}^* \boldsymbol{y}| \leq \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q$$
- Cauchy-Schwarz inequality
$$|\boldsymbol{x}^* \boldsymbol{y}| \leq \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2$$
- Cauchy-Schwarz inequality is a special case of Hölder inequality
- Example: What is 2-norm of rank-one matrix? Hint: Use Cauchy-Schwarz inequality.

Bounding Matrix-Matrix Multiplication

- Let $\boldsymbol{A}$ be an $l \times m$ matrix and $\boldsymbol{B}$ an $m \times n$ matrix, then for $\boldsymbol{x} \in \mathbb{C}^n$

$$\|\boldsymbol{AB}\|_{(l,n)} \leq \|\boldsymbol{A}\|_{(l,m)} \|\boldsymbol{B}\|_{(m,n)}$$

Bounding Matrix-Matrix Multiplication

- Let $\boldsymbol{A}$ be an $l \times m$ matrix and $\boldsymbol{B}$ an $m \times n$ matrix, then for $\boldsymbol{x} \in \mathbb{C}^n$

$$\|\boldsymbol{AB}\|_{(l,n)} \leq \|\boldsymbol{A}\|_{(l,m)}\|\boldsymbol{B}\|_{(m,n)}$$

- To show it, note

$$\|\boldsymbol{ABx}\|_{(l)} \leq \|\boldsymbol{A}\|_{(l,m)}\|\boldsymbol{Bx}\|_{(m)} \leq \|\boldsymbol{A}\|_{(l,m)}\|\boldsymbol{B}\|_{(m,n)}\|\boldsymbol{x}\|_{(n)},$$

Bounding Matrix-Matrix Multiplication

- Let $\boldsymbol{A}$ be an $l \times m$ matrix and $\boldsymbol{B}$ an $m \times n$ matrix, then for $\boldsymbol{x} \in \mathbb{C}^n$

$$\|\boldsymbol{AB}\|_{(l,n)} \leq \|\boldsymbol{A}\|_{(l,m)}\|\boldsymbol{B}\|_{(m,n)}$$

- To show it, note

$$\|\boldsymbol{ABx}\|_{(l)} \leq \|\boldsymbol{A}\|_{(l,m)}\|\boldsymbol{Bx}\|_{(m)} \leq \|\boldsymbol{A}\|_{(l,m)}\|\boldsymbol{B}\|_{(m,n)}\|\boldsymbol{x}\|_{(n)},$$

- In general, this inequality is not an equality
- In particular, $\|\boldsymbol{A}^n\| \leq \|\boldsymbol{A}\|^n$ but $\|\boldsymbol{A}^n\| \neq \|\boldsymbol{A}\|^n$ in general for $n \geq 2$

Bounding Matrix-Matrix Multiplication

- Let $\boldsymbol{A}$ be an $l \times m$ matrix and $\boldsymbol{B}$ an $m \times n$ matrix, then for $\boldsymbol{x} \in \mathbb{C}^n$

$$\|\boldsymbol{AB}\|_{(l,n)} \leq \|\boldsymbol{A}\|_{(l,m)} \|\boldsymbol{B}\|_{(m,n)}$$

- To show it, note

$$\|\boldsymbol{ABx}\|_{(l)} \leq \|\boldsymbol{A}\|_{(l,m)} \|\boldsymbol{Bx}\|_{(m)} \leq \|\boldsymbol{A}\|_{(l,m)} \|\boldsymbol{B}\|_{(m,n)} \|\boldsymbol{x}\|_{(n)},$$

- In general, this inequality is not an equality
- In particular, $\|\boldsymbol{A}^n\| \leq \|\boldsymbol{A}\|^n$ but $\|\boldsymbol{A}^n\| \neq \|\boldsymbol{A}\|^n$ in general for $n \geq 2$
- Also, note that

$$\|\boldsymbol{AB}\|_F \leq \|\boldsymbol{A}\|_F \|\boldsymbol{B}\|_F$$

because

$$\|\boldsymbol{AB}\|_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} |\boldsymbol{a}_i^* \boldsymbol{b}_j|^2 \leq \sum_{i=1}^{n} \sum_{j=1}^{m} (\|\boldsymbol{a}_i^*\|_2 \|\boldsymbol{b}_j\|_2)^2 = \|\boldsymbol{A}\|_F^2 \|\boldsymbol{B}\|_F^2$$