NE 255, Class 3, Fa16
Vectors and Matrix norms, Convergence
September 01, 2016

Introduction

In today's lecture we will introduce/recap basic concepts about vector and matrix norms.

- Vector norms
  - l_1, l_2, and l_∞
  - Triangle inequality and Cauchy-Schwarz
  - Convergence
  - Equivalence of norms
- Matrix norms
  - Natural norms
  - Convergence
  - Spectral Radius
  - Gelfand's Formula

1 Vector Norms

Definition A *Vector Norm* on $\mathbb{R}^n$ is a function $||\cdot||$ mapping $\mathbb{R}^n \to \mathbb{R}$ with the following properties:

1. $||\vec{x}|| \geq 0$ for all $\vec{x} \in \mathbb{R}^n$
2. $||\vec{x}|| = 0$ iff $\vec{x} = 0$
3. $||\alpha\vec{x}|| = |\alpha|||\vec{x}||$ for all $\alpha \in \mathbb{R}$ and $\vec{x} \in \mathbb{R}^n$ (scalar multiplication)
4. $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$ (triangle inequality)

Common Norms

- The l_1 norm is given by

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

- The *Max norm, Sup norm*, or l_∞ norm, is given by

$$||\vec{x}||_\infty = \max_{1\leq i\leq n} |x_i|$$

- The *Euclidean Norm*, or l_2 norm, is given by

$$||\vec{x}||_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$$

  Note: this norm represents the usual notion of distance (Pythagorean theorem)

Example Consider $\vec{x} = (1, -3, 5)^T$. Then

i. $||\vec{x}||_1 = 9$

ii. $||\vec{x}||_\infty = 5$

iii. $||\vec{x}||_2 = \sqrt{1+9+25} = \sqrt{35}$

$- - - - - - - - - - - - - - - - - - - -$

Cauchy-Schwarz and Triangle Inequality

It is easy to see that all the properties of a vector norm are satisfied for l_1 and l_∞, but we need to show that the triangle inequality holds for l_2. We will need the following theorem:

Theorem 1. *(Cauchy-Schwarz in $\mathbb{R}^n$)*
For each $\vec{x}, \vec{y} \in \mathbb{R}^n$,

$$\vec{x}^T\vec{y} = \sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \left(\sum_{i=1}^{n} y_i^2\right)^{1/2} = ||\vec{x}||_2||\vec{y}||_2$$

Proof. Consider the following quadratic polynomial in $z \in \mathbb{R}$:

$$0 \leq (x_1 z + y_1)^2 + ... + (x_n z + y_n)^2 = \left(\sum_{i=1}^{n} x_i^2\right) z^2 + 2\left(\sum_{i=1}^{n} x_i y_i\right) z + \left(\sum_{i=1}^{n} y_i^2\right).$$

Since it is nonnegative, it has at most one real root for z. Hence, its discriminant is less than or equal to zero; that is,

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 - \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) \leq 0.$$

□

Proof of the Triangle Inequality for l_2:

Proof. Using Cauchy-Schwarz, for each $\vec{x}, \vec{y} \in \mathbb{R}^n$:

$$||\vec{x}+\vec{y}||_2^2 = \sum_{i=1}^{n}(x_i + y_i)^2 = \sum_{i=1}^{n} x_i^2 + 2\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} y_i^2 \leq ||\vec{x}||_2^2 + 2||\vec{x}||_2||\vec{y}||_2 + ||\vec{y}||_2^2.$$

Taking the square root of both sides, we obtain

$$||\vec{x}+\vec{y}||_2 \leq ||\vec{x}||_2 + ||\vec{y}||_2.$$

□

Basic Theorems of Convergence

Definition A sequence of vectors $\{\vec{x}^{(k)}\}_{k=1}^{\infty}$ in $\mathbb{R}^n$ is said to *converge* to $\vec{x}$ with respect to norm $||\cdot||$ if given any $\varepsilon > 0$ there exists an integer $N(\varepsilon)$ such that

$$||\vec{x}^{(k)} - \vec{x}|| < \varepsilon \quad \text{for all} \quad k \geq N(\varepsilon).$$

Theorem 2. *The sequence of vectors $\{\vec{x}^{(k)}\}_{k=1}^{\infty} \to \vec{x}$ in $\mathbb{R}^n$ with respect to $||\cdot||_\infty$ iff*

$$\lim_{k\to\infty} x_i^{(k)} = x_i \quad \textit{for each} \quad i = 1, 2, ..., n.$$

Proof. ($\hookrightarrow$)

Let $\lim_{k\to\infty} ||\vec{x}^{(k)}||_\infty = ||\vec{x}||_\infty$. Then for any $\varepsilon > 0$, there exists $N(\varepsilon)$ such that

$$||\vec{x}^{(j)} - \vec{x}^{(m)}||_\infty < \varepsilon \quad \text{for all} \quad j, m > N(\varepsilon).$$

Thus

$$\max_{1\le i\le n} |x_i^{(j)} - x_i^{(m)}| < \varepsilon \quad \text{for all} \quad j, m > N(\varepsilon),$$

implying that

$$|x_i^{(j)} - x_i^{(m)}| < \varepsilon \quad \text{for all} \quad i \quad \text{and for all} \quad j, m > N(\varepsilon).$$

Therefore,

$$\lim_{k\to\infty} x_i^{(k)} = x_i \quad \text{for each} \quad i = 1, 2, ..., n.$$

□

Proof. ($\hookleftarrow$)

Let

$$\lim_{k\to\infty} x_i^{(k)} = x_i \quad \text{for each} \quad i = 1, 2, ..., n.$$

Then for any $\varepsilon > 0$, there exists $N(\varepsilon)$ such that

$$|x_i^{(j)} - x_i^{(m)}| < \frac{\varepsilon}{2} \quad \text{for all} \quad i \quad \text{and for all} \quad j, m > N(\varepsilon).$$

Taking the limit as $m \to \infty$, we have

$$|x_i^{(j)} - x_i| \le \frac{\varepsilon}{2} \quad \text{for all} \quad i \quad \text{and for all} \quad j > N(\varepsilon).$$

Thus,

$$\max_{1\le i\le n} |x_i^{(j)} - x_i| \le \frac{\varepsilon}{2} \quad \text{for all} \quad j > N(\varepsilon),$$

which means that

$$\lim_{j\to\infty} ||\vec{x}^{(j)} - \vec{x}||_\infty \le \frac{\varepsilon}{2} < \varepsilon.$$

□

Theorem 3. *For each $\vec{x} \in \mathbb{R}^n$:*

a. $||\vec{x}||_\infty \le ||\vec{x}||_2 \le \sqrt{n}||\vec{x}||_\infty$

b. $||\vec{x}||_2 \le ||\vec{x}||_1 \le \sqrt{n}||\vec{x}||_2$

c. $||\vec{x}||_\infty \leq ||\vec{x}||_1 \leq n||\vec{x}||_\infty$

Proof. We give the proof for $(a.)$:

$$||\vec{x}||_2 = ||\vec{x}||_\infty \left(\sum_{i=1}^{n} \frac{x_i^2}{||\vec{x}||_\infty^2} \right)^{1/2} \leq ||\vec{x}||_\infty \sqrt{n},$$

because $x_i/||\vec{x}||_\infty \leq 1$ for all i.

Moreover, there is a i such that $||\vec{x}||_\infty = |x_i|$, therefore

$$\left(\sum_{i=1}^{n} \frac{x_i^2}{||\vec{x}||_\infty^2} \right)^{1/2} \geq 1$$

and

$$||\vec{x}||_2 = ||\vec{x}||_\infty \left(\sum_{i=1}^{n} \frac{x_i^2}{||\vec{x}||_\infty^2} \right)^{1/2} \geq ||\vec{x}||_\infty.$$

□

Note: As a corollary of this theorem, convergence in the l_1, l_2, and l_∞ norms is equivalent.

2 Matrix Norms

We need to extend our definitions to include matrices.

Definition A *Matrix Norm* on the set of all $n \times n$ matrices is a real-valued function $|| \cdot ||$ defined on this set that satisfies the following properties for all $n \times n$ matrices A and B and all real numbers α:

1. $||A|| \geq 0$
2. $||A|| = 0$ iff $A = 0$ (all zero entries)
3. $||\alpha A|| = |\alpha| ||A||$ (scalar multiplication)
4. $||A + B|| \leq ||A|| + ||B||$ (triangle inequality)

In this course, in the case of square matrices, we will deal with submultiplicative norms, which also satisfy

5. $||AB|| \leq ||A||\,||B||$

The following theorem is offered without proof:

Theorem 4. *(Natural or Induced Matrix Norm)*
If $||\cdot||$ is a vector norm on $\mathbb{R}^n$, then

$$||A|| = \max_{||\vec{x}||=1} ||A\vec{x}||$$

is a matrix norm.

The natural norm describes how a matrix stretches unit vectors relative to that norm. For any $\vec{y} \neq 0$, $\vec{x} = \vec{y}/||\vec{y}||$ is a unit vector, and

$$\max_{||\vec{x}||=1} ||Ax|| = \max_{||\vec{y}||\neq 0} \left|\left| A\frac{\vec{y}}{||\vec{y}||} \right|\right| = \max_{||\vec{y}||\neq 0} \frac{||A\vec{y}||}{||\vec{y}||}.$$

<u>**Common Norms**</u>

- $||A||_1 = \max\limits_{1\leq j\leq n} \sum\limits_{i=1}^{m} |a_{ij}|$ = largest absolute column sum.
- $||A||_\infty = \max\limits_{1\leq i\leq m} \sum\limits_{i=1}^{n} |a_{ij}|$ = largest absolute row sum.
- In the special case of the Euclidean norm, the induced matrix norm is the *Spectral Norm*. The spectral norm of a matrix A is the largest singular value of A; i.e. the square root of the largest eigenvalue of the positive-semidefinite matrix A^*A:

  $$||A||_2 = \sqrt{\lambda_{max}(A^*A)} = \sigma_{max}(A).$$

  It can be shown that

  $$||A||_2 \leq \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2} = ||A||_F,$$

  where the right-hand side is the Frobenius norm, or $L_{2,2}$ norm. The equality holds if and only if the matrix A is a rank-one matrix or a zero matrix.

Example Consider

$$A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}.$$

Then

i. $||A||_1 = 3$

ii. $||A||_\infty = 2$

iii. $||A||_2 = \sqrt{3+\sqrt{5}} \approx 2.2882$

$- - - - - - - - - - - - - - - - - -$

<u>**Convergence and Spectral Radius**</u>

Definition An $n \times n$ matrix A is convergent if

$$\lim_{k\to\infty} (A^k)_{ij} = 0,$$

for each $i, j = 1, 2, ..., n$.

Example Consider

$$A = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

We can see that

$$A^k = \begin{pmatrix} \frac{1}{2^k} & 0 \\ \frac{k}{2^{k+1}} & \frac{1}{2^k} \end{pmatrix} \to 0$$

as $k \to \infty$.

$- - - - - - - - - - - - - - - - - -$

Definition The spectral radius, $\rho(A)$, of a matrix A is defined by

$$\rho(A) = \max |\lambda|,$$

where λ is an eigenvalue of A.

The spectral radius provides a valuable measure of the eigenvalues, which helps determine if a numerical scheme will converge.

Theorem 5. *If $A \in \mathbb{R}^{n\times n}$, then $\rho(A) \leq ||A||$ for any natural norm $||\cdot||$.*

Proof. Let $||\vec{x}||$ be a unit eigenvector of A with respect to the eigenvalue λ. Then

$$|\lambda| = |\lambda|\,||\vec{x}|| = ||\lambda\vec{x}|| = ||Ax|| \leq ||A||\,||x|| = ||A||.$$

Thus,

$$\rho(A) = \max|\lambda| \leq ||A||.$$

If A is symmetric, then $\rho(A) = ||A||_2$. □

The following theorem is given here without proof. It will be used to prove Gelfand's Formula.

Theorem 6. *Let $A \in \mathbb{C}^{n\times n}$ with spectral radius $\rho(A)$; then $\rho(A) < 1$ iff*

$$\lim_{k\to\infty} A^k = 0.$$

Moreover, if $\rho(A) > 1$, $||A^k||$ is not bounded for increasing values of k.

Theorem 7. *(Gelfand's Formula)*

For any matrix norm $||\cdot||$, we have

$$\rho(A) = \lim_{k\to\infty} ||A^k||^{1/k}.$$

Proof. For any $\varepsilon > 0$ we construct the following two matrices:

$$A_{\pm} = \frac{1}{\rho(A) \pm \varepsilon} A.$$

Then

$$\rho(A_{\pm}) = \frac{\rho(A)}{\rho(A) \pm \varepsilon},$$

which implies that $\rho(A_+) < 1 < \rho(A_-)$.

Using Theorem **6**, there exists $N_+ \in \mathbb{N}$ such that $\forall\ k \geq N_+$, we have $||A_+^k|| < 1$ and therefore

$$\forall\ k \geq N_+ \quad ||A^k|| < (\rho(A) + \varepsilon)^k,$$

and

$$\forall\ k \geq N_+ \quad ||A^k||^{1/k} < \rho(A) + \varepsilon.$$

Applying Theorem **6** to A_- implies that $||A_-^k||$ is not bounded and there exists $N_- \in \mathbb{N}$ such that $\forall\ k \geq N_-$ we have $||A_-^k|| > 1$ and therefore

$$\forall\ k \geq N_- \quad ||A^k|| > (\rho(A) - \varepsilon)^k,$$

and

$$\forall\ k \geq N_- \quad ||A^k||^{1/k} > \rho(A) - \varepsilon.$$

Let $N = max\{N_+, N_-\}$, then we have that $\forall\ \varepsilon > 0$, there is $N \in \mathbb{N}$ such that $\forall\ k \geq N$

$$\rho(A) - \varepsilon < ||A^k||^{1/k} < \rho(A) + \varepsilon$$

which implies that

$$\lim_{k\to\infty} ||A^k||^{1/k} = \rho(A).$$

□

Finally, we have the following result, offered without proof:

Theorem 8. *The following statements are equivalent:*

a. A *is a convergent matrix*

b. $\lim\limits_{n\to\infty} ||A^n|| = 0$ *for some natural norm*

c. $\lim\limits_{n\to\infty} ||A^n|| = 0$ *for all natural norms*

d. $\rho(A) < 1$

e. $\lim\limits_{n\to\infty} A^n\vec{x} = 0$ *for every* $\vec{x}$