Numerical Analysis

NTNU

Tsung-Min Hwang

September 11, 2003

1 Norms

1.1 Vector Norm Definition and Properties

Definition 1.1 *A vector norm is a function* $\|\cdot\| : \mathbb{R}^n \longrightarrow \mathbb{R}$ *satisfying the following conditions for all* $x, y \in \mathbb{R}^n$ *and* $\alpha \in \mathbb{R}$*.*

1. $\|x\| \geq 0 \quad (\|x\| = 0 \Leftrightarrow x = 0)$;
2. $\|x + y\| \leq \|x\| + \|y\|$;
3. $\|\alpha x\| = |\alpha|\|x\|$.

Some of the most frequently used vector norms for $x \in \mathbb{R}^n$:

☞ 1**-norm:**

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|.$$

☞ 1**-norm:**

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|.$$

☞ 2**-norm:**

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \left(|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2\right)^{1/2} = \sqrt{x^T x}.$$

☞ 1**-norm:**

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|.$$

☞ 2**-norm:**

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \left(|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2\right)^{1/2} = \sqrt{x^T x}.$$

☞ ∞**-norm:**

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|.$$

☞ 1-**norm:**

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|.$$

☞ 2-**norm:**

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \left(|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2\right)^{1/2} = \sqrt{x^T x}.$$

☞ ∞-**norm:**

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|.$$

☞ p-**norm:**

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} = (|x_1|^p + |x_2|^p + \ldots + |x_n|^p)^{1/p}.$$

Definition 1.2 (unit vector) *$x \in \mathbb{R}^n$ is called a unit vector if $\|x\| = 1$ with respect to some vector norm.*

Definition 1.2 (unit vector) $x \in \mathbb{R}^n$ *is called a unit vector if* $\|x\| = 1$ *with respect to some vector norm.*

Property 1.1 *For any* $x, y \in \mathbb{R}^n$, *the following two inequalities hold.*

- *Hölder inequality:*

$$|x^T y| \leq \|x\|_p \|y\|_q, \quad \text{where} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

- *Cauchy-Schwartz inequality:*

$$|x^T y| \leq \|x\|_2 \|y\|_2.$$

Definition 1.2 (unit vector) $x \in \mathbb{R}^n$ *is called a unit vector if* $\|x\| = 1$ *with respect to some vector norm.*

Property 1.1 *For any* $x, y \in \mathbb{R}^n$*, the following two inequalities hold.*

- *Hölder inequality:*

$$|x^T y| \le \|x\|_p \|y\|_q, \quad \text{where} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

- *Cauchy-Schwartz inequality:*

$$|x^T y| \le \|x\|_2 \|y\|_2.$$

Definition 1.3 *Two vector norms* $\|\cdot\|_\alpha$ *and* $\|\cdot\|_\beta$ *are equivalent if there exist constants* $c_1, c_2 \in \mathbb{R}$ *such that*

$$c_1 \|x\|_\alpha \le \|x\|_\beta \le c_2 \|x\|_\alpha$$

for any $x \in \mathbb{R}^n$*.*

Property 1.2 *For all* $x \in \mathbb{R}^n$,

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2. \quad (1)$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty. \quad (2)$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty. \quad (3)$$

Property 1.2 *For all $x \in \mathbb{R}^n$,*

$$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2. \tag{1}$$

$$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty. \tag{2}$$

$$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty. \tag{3}$$

Definition 1.4 (absolute error and relative error) *Suppose $x \in \mathbb{R}^n$ is the exact solution of some problem and $\tilde{x}$ is an approximation to x. We define*

$$\textit{absolute error} = \|x - \tilde{x}\| \tag{4}$$

and

$$\textit{relative error} = \frac{\|x - \tilde{x}\|}{\|x\|}, \quad \textit{if} \quad x \neq 0. \tag{5}$$

1.2 Matrix Norm Definition and Properties

Definition 1.5 *A matrix norm is a function* $\| \cdot \| : \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ *satisfying the following conditions for all* $A, B \in \mathbb{R}^{m \times n}$ *and* $\alpha \in \mathbb{R}$*.*

1. $\|A\| \geq 0 \quad (\|A\| = 0 \Leftrightarrow A = 0)$;
2. $\|A + B\| \leq \|A\| + \|B\|$;
3. $\|\alpha A\| = |\alpha| \|A\|$.

Definition 1.6 *Some of the most frequently used matrix norms are*

1.2 Matrix Norm Definition and Properties

Definition 1.5 *A matrix norm is a function* $\|\cdot\| : \mathbb{R}^{m\times n} \rightarrow \mathbb{R}$ *satisfying the following conditions for all* $A, B \in \mathbb{R}^{m\times n}$ *and* $\alpha \in \mathbb{R}$*.*

1. $\|A\| \geq 0 \quad (\|A\| = 0 \Leftrightarrow A = 0)$;
2. $\|A + B\| \leq \|A\| + \|B\|$;
3. $\|\alpha A\| = |\alpha|\|A\|$.

Definition 1.6 *Some of the most frequently used matrix norms are*

☞ **Frobenius norm:**

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}.$$

☞ 2**-norm:**

$$\|A\|_2 = \max_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2} = \max_{\|x\|_2 = 1} \|Ax\|_2.$$

☞ 2-**norm:**

$$\|A\|_2 = \max_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2} = \max_{\|x\|_2 = 1} \|Ax\|_2.$$

☞ 1-**norm:**

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

☞ 2**-norm:**

$$\|A\|_2 = \max_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2} = \max_{\|x\|_2 = 1} \|Ax\|_2.$$

☞ 1**-norm:**

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

☞ ∞**-norm:**

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

☞ 2-**norm:**

$$\|A\|_2 = \max_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2} = \max_{\|x\|_2 = 1} \|Ax\|_2.$$

☞ 1-**norm:**

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|.$$

☞ ∞-**norm:**

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|.$$

☞ p-**norm:**

$$\|A\|_p = \max_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|_p}{\|x\|_p} = \max_{\|x\|_p = 1} \|Ax\|_p.$$

Theorem 1.1 *Suppose* $A \in \mathbb{R}^{m\times n}$*. Then there exists* $z \in \mathbb{R}^n$*,* $\|z\|_2 = 1$*, such that* $A^T A z = \mu^2 z$*, where* $\mu = \|A\|_2$*.*

Proof: Let $z \in \mathbb{R}^n$, $\|z\|_2 = 1$, be a unit vector that satisfies $\|A\|_2 = \|Az\|_2 = \max_{\|x\|_2=1} \|Ax\|_2$. Define

$$g(x) = \frac{1}{2}\frac{x^T A^T A x}{x^T x} = \frac{1}{2}\frac{\|Ax\|_2^2}{\|x\|_2^2} = \frac{1}{2}\left(\frac{\|Ax\|_2}{\|x\|_2}\right)^2,$$

for $x \in \mathbb{R}^n$. Then z is a maximizer of $g(x)$ which implies $\bigtriangledown g(z) = 0$. Since

$$\bigtriangledown g(x) = \frac{(x^T x)(A^T A x) - (x^T A^T A x)(x)}{(x^T x)^2}.$$

Hence

$$\begin{aligned}
\bigtriangledown g(z) = 0 \quad &\Rightarrow \quad (z^T z)(A^T A z) - (z^T A^T A z) z = 0 \\
&\Rightarrow \quad \|z\|_2^2 (A^T A z) - \|Az\|_2^2 z = 0 \\
&\Rightarrow \quad A^T A z = \|A\|_2^2 z = \mu^2 z.
\end{aligned}$$

■

Remarks 1.1 *$\|A\|_2$ is the square root of the largest eigenvalue of $A^T A$. When A is symmetric, $\|A\|_2$ is the absolute value of the largest eigenvalue in magnitude.*

Remarks 1.1 *$\|A\|_2$ is the square root of the largest eigenvalue of $A^T A$. When A is symmetric, $\|A\|_2$ is the absolute value of the largest eigenvalue in magnitude.*

Property 1.3

$$\|A\|_2 \le \|A\|_F \le \sqrt{n}\|A\|_2. \tag{6}$$

$$\frac{1}{\sqrt{n}}\|A\|_\infty \le \|A\|_2 \le \sqrt{m}\|A\|_\infty. \tag{7}$$

$$\frac{1}{\sqrt{m}}\|A\|_1 \le \|A\|_2 \le \sqrt{n}\|A\|_1. \tag{8}$$

$$\max_{i,j} |a_{ij}| \le \|A\|_2 \le \sqrt{mn}\max_{i,j} |a_{ij}|. \tag{9}$$

☞ SVD: The Singular Value Decomposition

☞ SVD: The Singular Value Decomposition

Theorem 1.2 (Existence of SVD) *If $A \in \mathbb{R}^{m \times n}$, then there exists orthogonal matrices $U = [u_1, u_2, \ldots, u_m] \in \mathbb{R}^{m \times m}$ and $V = [v_1.v_2, \ldots, v_n] \in \mathbb{R}^{n \times n}$ such that*

$$A = U \Sigma V^T, \tag{10}$$

where

$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_p), \qquad p = \min(m, n),$$

with

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0.$$

☞ SVD: The Singular Value Decomposition

Theorem 1.2 (Existence of SVD) *If $A \in \mathbb{R}^{m\times n}$, then there exists orthogonal matrices $U = [u_1, u_2, \ldots, u_m] \in \mathbb{R}^{m\times m}$ and $V = [v_1.v_2, \ldots, v_n] \in \mathbb{R}^{n\times n}$ such that*

$$A = U\Sigma V^T, \tag{10}$$

where

$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_p), \qquad p = \min(m, n),$$

with

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0.$$

☞ The σ_i are called the singular values of A and the vectors u_i and v_i the i-th left singular vector and the i-th right singular vector, respectively.

☞ SVD: The Singular Value Decomposition

Theorem 1.2 (Existence of SVD) *If $A \in \mathbb{R}^{m\times n}$, then there exists orthogonal matrices $U = [u_1, u_2, \ldots, u_m] \in \mathbb{R}^{m\times m}$ and $V = [v_1.v_2, \ldots, v_n] \in \mathbb{R}^{n\times n}$ such that*

$$A = U\Sigma V^T, \tag{10}$$

where

$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_p), \qquad p = \min(m, n),$$

with

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0.$$

☞ The σ_i are called the singular values of A and the vectors u_i and v_i the i-th left singular vector and the i-th right singular vector, respectively.

☞ We usually use $\sigma_{max}(A)$ to denote the largest singular value of A and $\sigma_{min}(A)$ the smallest singular value of A.