MA423 Matrix Computations

Lecture 10: Vector and matrix norms

Rafikul Alam
Department of Mathematics
IIT Guwahati

Outline

- Vector norms
- Matrix norms
- Unitarily invariant norms

Vector norms

Let $\mathcal{V}$ be a vector space over $\mathbb{C}$. Then a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ is called a norm on $\mathcal{V}$ if it satisfies the three fundamental properties:

(a) Positive definiteness: $\|v\| \geq 0$ and $\|v\| = 0 \iff v = 0$.

(b) Positively homogeneous: $\|\alpha v\| = |\alpha| \, \|v\|$ for $\alpha \in \mathbb{C}$ and $v \in \mathcal{V}$.

(c) Triangle inequality: $\|u + v\| \leq \|u\| + \|v\|$ for $u, v \in \mathcal{V}$.

Vector norms

Let $\mathcal{V}$ be a vector space over $\mathbb{C}$. Then a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ is called a norm on $\mathcal{V}$ if it satisfies the three fundamental properties:

(a) Positive definiteness: $\|v\| \geq 0$ and $\|v\| = 0 \iff v = 0$.

(b) Positively homogeneous: $\|\alpha v\| = |\alpha| \, \|v\|$ for $\alpha \in \mathbb{C}$ and $v \in \mathcal{V}$.

(c) Triangle inequality: $\|u + v\| \leq \|u\| + \|v\|$ for $u, v \in \mathcal{V}$.

Example: Consider $\mathbb{C}^n$ and the vector norms given by

1-norm: $\|x\|_1 := |x_1| + \cdots + |x_n|$.

2-norm: $\|x\|_2 := \sqrt{|x_1|^2 + \cdots |x_n|^2}$.

∞-norm: $\|x\|_\infty := \max_{1 \leq j \leq n} |x_j|$.

Vector norms

Let $\mathcal{V}$ be a vector space over $\mathbb{C}$. Then a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ is called a norm on $\mathcal{V}$ if it satisfies the three fundamental properties:

(a) Positive definiteness: $\|v\| \geq 0$ and $\|v\| = 0 \iff v = 0$.

(b) Positively homogeneous: $\|\alpha v\| = |\alpha| \, \|v\|$ for $\alpha \in \mathbb{C}$ and $v \in \mathcal{V}$.

(c) Triangle inequality: $\|u + v\| \leq \|u\| + \|v\|$ for $u, v \in \mathcal{V}$.

Example: Consider $\mathbb{C}^n$ and the vector norms given by

1-norm: $\|x\|_1 := |x_1| + \cdots + |x_n|$.

2-norm: $\|x\|_2 := \sqrt{|x_1|^2 + \cdots |x_n|^2}$.

∞-norm: $\|x\|_\infty := \max_{1 \leq j \leq n} |x_j|$.

More generally, for $1 \leq p < \infty$, the Höder p-norm is given by

$$\|x\|_p := (|x_1|^p + \cdots + |x_n|^p)^{1/p}.$$

Matrix norms

Let $A \in \mathbb{C}^{m\times n}$. Then $A : \mathbb{C}^n \longrightarrow \mathbb{C}^m$, $x \longmapsto Ax$, is a linear map. Suppose $\mathbb{C}^n$ and $\mathbb{C}^m$ are equipped with norms. Then

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

defines a norm on $\mathbb{C}^{m\times n}$ and is called an induced matrix norm or a subordinate matrix norm.

Matrix norms

Let $A \in \mathbb{C}^{m\times n}$. Then $A : \mathbb{C}^n \longrightarrow \mathbb{C}^m$, $x \longmapsto Ax$, is a linear map. Suppose $\mathbb{C}^n$ and $\mathbb{C}^m$ are equipped with norms. Then

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

defines a norm on $\mathbb{C}^{m\times n}$ and is called an induced matrix norm or a subordinate matrix norm.

For the identity matrix $\|Ix\| = \|x\|$ and hence $\|I\| = 1$. Note that

$$\|Ax\| \leq \|A\| \, \|x\|$$

for all $x \in \mathbb{C}^n$.

Matrix norms

The matrix norm of A induced by the Hölder p-norm is denoted by $\|A\|_p$. Then $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$ are called 1-norm, 2-norm and ∞-norm of A, respectively. Also $\|A\|_2$ is called the spectral norm of A.

Matrix norms

The matrix norm of A induced by the Hölder p-norm is denoted by $\|A\|_p$. Then $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$ are called 1-norm, 2-norm and ∞-norm of A, respectively. Also $\|A\|_2$ is called the spectral norm of A.

Theorem: Let A be an $m \times n$ matrix. Then

$$\begin{aligned} \|A\|_1 &= \max_{1 \le j \le n} \|Ae_j\|_1 = \max_{1 \le j \le n} \|A(:,j)\|_1 \\ \|A\|_2 &= \sqrt{\lambda_{\max}(A^*A)} \\ \|A\|_\infty &= \max_{1 \le i \le m} \|e_i^\top A\|_1 = \max_{1 \le i \le m} \|A(i,:)\|_1, \end{aligned}$$

where $\lambda_{\max}(A^*A)$ denotes that largest eigenvalue of A^*A.

Matrix norms

The matrix norm of A induced by the Hölder p-norm is denoted by $\|A\|_p$. Then $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$ are called 1-norm, 2-norm and ∞-norm of A, respectively. Also $\|A\|_2$ is called the spectral norm of A.

Theorem: Let A be an $m \times n$ matrix. Then

$$\begin{aligned} \|A\|_1 &= \max_{1\leq j\leq n} \|Ae_j\|_1 = \max_{1\leq j\leq n} \|A(:,j)\|_1 \\ \|A\|_2 &= \sqrt{\lambda_{\max}(A^*A)} \\ \|A\|_\infty &= \max_{1\leq i\leq m} \|e_i^\top A\|_1 = \max_{1\leq i\leq m} \|A(i,:)\|_1, \end{aligned}$$

where $\lambda_{\max}(A^*A)$ denotes that largest eigenvalue of A^*A.

Proof: We have
$Ax = x_1Ae_1 + \cdots + x_nAe_n \Rightarrow \|Ax\|_1 \leq \max_{1\leq j\leq n} \|Ae_j\|_1\|x\|_1$.

Matrix norms

The matrix norm of A induced by the Hölder p-norm is denoted by $\|A\|_p$. Then $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$ are called 1-norm, 2-norm and ∞-norm of A, respectively. Also $\|A\|_2$ is called the spectral norm of A.

Theorem: Let A be an $m \times n$ matrix. Then

$$\begin{aligned}
\|A\|_1 &= \max_{1\leq j\leq n} \|Ae_j\|_1 = \max_{1\leq j\leq n} \|A(:,j)\|_1 \\
\|A\|_2 &= \sqrt{\lambda_{\max}(A^*A)} \\
\|A\|_\infty &= \max_{1\leq i\leq m} \|e_i^\top A\|_1 = \max_{1\leq i\leq m} \|A(i,:)\|_1,
\end{aligned}$$

where $\lambda_{\max}(A^*A)$ denotes that largest eigenvalue of A^*A.

Proof: We have
$Ax = x_1Ae_1 + \cdots + x_nAe_n \Rightarrow \|Ax\|_1 \leq \max_{1\leq j\leq n} \|Ae_j\|_1 \|x\|_1$. This yields $\|A\|_1 \leq \max_{1\leq j\leq n} \|Ae_j\|_1$.

Matrix norms

The matrix norm of A induced by the Hölder p-norm is denoted by $\|A\|_p$. Then $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$ are called 1-norm, 2-norm and ∞-norm of A, respectively. Also $\|A\|_2$ is called the spectral norm of A.

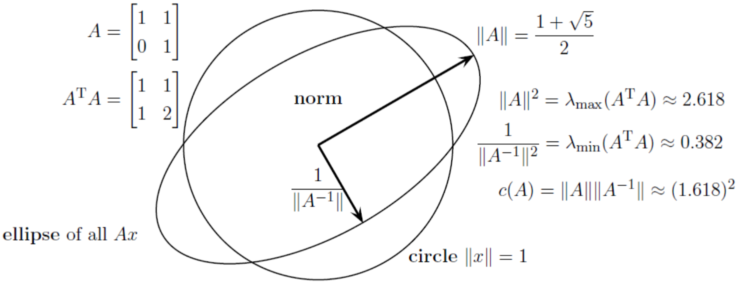
Theorem: Let A be an $m \times n$ matrix. Then

$$\begin{aligned} \|A\|_1 &= \max_{1\leq j\leq n} \|Ae_j\|_1 = \max_{1\leq j\leq n} \|A(:,j)\|_1 \\ \|A\|_2 &= \sqrt{\lambda_{\max}(A^*A)} \\ \|A\|_\infty &= \max_{1\leq i\leq m} \|e_i^\top A\|_1 = \max_{1\leq i\leq m} \|A(i,:)\|_1, \end{aligned}$$

where $\lambda_{\max}(A^*A)$ denotes that largest eigenvalue of A^*A.

Proof: We have $Ax = x_1Ae_1 + \cdots + x_nAe_n \Rightarrow \|Ax\|_1 \leq \max_{1\leq j\leq n} \|Ae_j\|_1\|x\|_1$. This yields $\|A\|_1 \leq \max_{1\leq j\leq n} \|Ae_j\|_1$. But $\|Ae_j\|_1 \leq \|A\|_1$ for all $j = 1 : n$. Hence we have $\|A\|_1 = \max_{1\leq j\leq n} \|Ae_j\|_1$. ■

Spectral norm



Frobenius norm

The spectral norm of a matrix is difficult to compute. By contrast, Frobenius norm is easy to compute and provides a nice geometric framework for solving problems. Let A be an m-by-n matrix.

Frobenius norm

The spectral norm of a matrix is difficult to compute. By contrast, Frobenius norm is easy to compute and provides a nice geometric framework for solving problems. Let A be an m-by-n matrix.

Then the Frobenius norm is given by

$$\|A\|_F := \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|A(i,j)|^2\right)^{1/2} = \sqrt{\mathrm{Tr}(A^*A)} = \sqrt{\mathrm{Tr}(AA^*)},$$

where $\mathrm{Tr}(A^*A)$ is the trace of A^*A.

Frobenius norm

The spectral norm of a matrix is difficult to compute. By contrast, Frobenius norm is easy to compute and provides a nice geometric framework for solving problems. Let A be an m-by-n matrix.

Then the Frobenius norm is given by

$$\|A\|_F := \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |A(i,j)|^2 \right)^{1/2} = \sqrt{\mathrm{Tr}(A^*A)} = \sqrt{\mathrm{Tr}(AA^*)},$$

where $\mathrm{Tr}(A^*A)$ is the trace of A^*A.

Note that $\|A\|_F = \left(\sum_{i=1}^{m} \|e_i^\top A\|_2^2\right)^{1/2} = \left(\sum_{j=1}^{n} \|Ae_j\|_2^2\right)^{1/2}$. Also note that for an n-by-n identity matrix I, we have $\|I\|_F = \sqrt{n}$ which shows that the Frobenius norm is not an induced matrix norm.

As induced matrix norm for Identity matrix is 1.

Example

Let $D := \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ be a diagonal matrix. Then we have $\|A\|_2 = \|A\|_1 = \|A\|_\infty = \max(|\lambda_1|, \ldots, |\lambda_n|)$ and $\|D\|_F = \sqrt{\sum_{j=1}^{n} |\lambda_j|^2}$.

Example

Let $D := \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ be a diagonal matrix. Then we have $\|A\|_2 = \|A\|_1 = \|A\|_\infty = \max(|\lambda_1|, \ldots, |\lambda_n|)$ and $\|D\|_F = \sqrt{\sum_{j=1}^n |\lambda_j|^2}$.

Consider $A := \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix}$. Then $A^{-1} = \begin{bmatrix} -998 & 999 \\ 999 & -1000 \end{bmatrix}$.

Example

Let $D := \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ be a diagonal matrix. Then we have $\|A\|_2 = \|A\|_1 = \|A\|_\infty = \max(|\lambda_1|, \ldots, |\lambda_n|)$ and $\|D\|_F = \sqrt{\sum_{j=1}^n |\lambda_j|^2}$.

Consider $A := \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix}$. Then $A^{-1} = \begin{bmatrix} -998 & 999 \\ 999 & -1000 \end{bmatrix}$.

Thus $\|A\|_\infty = \|A\|_1 = \|A^{-1}\|_\infty = \|A^{-1}\|_1 = 1999.$

On the other hand,

$$\|A\|_2 = 1998.000500500375 \text{ and } \|A\|_F = 1998.000500500438.$$

Submultiplicative matrix norm

A matrix norm is said to be submultiplicative if

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

holds for all A and B. An induced matrix norm is submultiplicative. Indeed, $\|ABx\| \leq \|A\| \, \|Bx\| \leq \|A\| \, \|B\| \, \|x\| \Longrightarrow \|AB\| \leq \|A\| \, \|B\|$.

Submultiplicative matrix norm

A matrix norm is said to be submultiplicative if

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

holds for all A and B. An induced matrix norm is submultiplicative. Indeed, $\|ABx\| \leq \|A\| \, \|Bx\| \leq \|A\| \, \|B\| \, \|x\| \Longrightarrow \|AB\| \leq \|A\| \, \|B\|$.

The Frobebnius norm is also submultiplicative. Indeed, suppose A is m-by-n and B is n-by-p matrices.

Submultiplicative matrix norm

A matrix norm is said to be submultiplicative if

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

holds for all A and B. An induced matrix norm is submultiplicative. Indeed, $\|ABx\| \leq \|A\| \, \|Bx\| \leq \|A\| \, \|B\| \, \|x\| \Longrightarrow \|AB\| \leq \|A\| \, \|B\|$.

The Frobebnius norm is also submultiplicative. Indeed, suppose A is m-by-n and B is n-by-p matrices. Then $(AB)_{ij} = e_i^\top ABe_j$ shows that $|(AB)_{ij}| \leq \|e_i^\top A\|_2 \|Be_j\|_2$ for $i = 1 : m$ and $j = 1 : n$. Consequently,

Submultiplicative matrix norm

A matrix norm is said to be submultiplicative if

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

holds for all A and B. An induced matrix norm is submultiplicative. Indeed, $\|ABx\| \leq \|A\| \, \|Bx\| \leq \|A\| \, \|B\| \, \|x\| \Longrightarrow \|AB\| \leq \|A\| \, \|B\|$.

The Frobebnius norm is also submultiplicative. Indeed, suppose A is m-by-n and B is n-by-p matrices. Then $(AB)_{ij} = e_i^\top ABe_j$ shows that $|(AB)_{ij}| \leq \|e_i^\top A\|_2 \|Be_j\|_2$ for $i = 1 : m$ and $j = 1 : n$. Consequently,

$$\begin{aligned}
\|AB\|_F^2 &= \sum_{i=1}^{m} \sum_{j=1}^{p} |(AB)_{ij}|^2 \leq \sum_{i=1}^{m} \sum_{j=1}^{p} \|e_i^\top A\|_2^2 \|Be_j\|_2^2 \\
&= \sum_{i=1}^{m} \|e_i^\top A\|_2^2 \sum_{j=1}^{p} \|Be_j\|_2^2 = \|A\|_F^2 \|B\|_F^2.
\end{aligned}$$

Unitarily invariant matrix norm

Let $Q \in \mathbb{R}^{n \times n}$. Then Q is said to be orthogonal if $Q^\top Q = QQ^\top = I$.

Unitarily invariant matrix norm

Let $Q \in \mathbb{R}^{n\times n}$. Then Q is said to be orthogonal if $Q^\top Q = QQ^\top = I$. If Q is orthogonal then
$\|Qx\|_2 = \sqrt{\langle Qx,\, Qx\rangle} = \sqrt{\langle x,\, Q^\top Qx\rangle} = \sqrt{\langle x,\, x\rangle} = \|x\|_2.$

Unitarily invariant matrix norm

Let $Q \in \mathbb{R}^{n\times n}$. Then Q is said to be orthogonal if $Q^\top Q = QQ^\top = I$. If Q is orthogonal then
$\|Qx\|_2 = \sqrt{\langle Qx,\, Qx\rangle} = \sqrt{\langle x,\, Q^\top Qx\rangle} = \sqrt{\langle x,\, x\rangle} = \|x\|_2.$

Let $Q \in \mathbb{C}^{n\times n}$. Then Q is said to be unitary if $Q^*Q = QQ^* = I$.

Unitarily invariant matrix norm

Let $Q \in \mathbb{R}^{n\times n}$. Then Q is said to be orthogonal if $Q^\top Q = QQ^\top = I$. If Q is orthogonal then
$\|Qx\|_2 = \sqrt{\langle Qx, Qx\rangle} = \sqrt{\langle x, Q^\top Qx\rangle} = \sqrt{\langle x, x\rangle} = \|x\|_2.$

Let $Q \in \mathbb{C}^{n\times n}$. Then Q is said to be unitary if $Q^* Q = QQ^* = I$. If Q is unitary then $\|Qx\|_2 = \sqrt{\langle Qx, Qx\rangle} = \sqrt{\langle x, Q^* Qx\rangle} = \sqrt{\langle x, x\rangle} = \|x\|_2.$

Unitarily invariant matrix norm

Let $Q \in \mathbb{R}^{n \times n}$. Then Q is said to be orthogonal if $Q^\top Q = QQ^\top = I$. If Q is orthogonal then
$\|Qx\|_2 = \sqrt{\langle Qx, Qx\rangle} = \sqrt{\langle x, Q^\top Qx\rangle} = \sqrt{\langle x, x\rangle} = \|x\|_2.$

Let $Q \in \mathbb{C}^{n \times n}$. Then Q is said to be unitary if $Q^* Q = QQ^* = I$. If Q is unitary then $\|Qx\|_2 = \sqrt{\langle Qx, Qx\rangle} = \sqrt{\langle x, Q^* Qx\rangle} = \sqrt{\langle x, x\rangle} = \|x\|_2.$

It follows that $\|Q\|_2 = 1$ when Q is unitary and that error in x is NOT magnified in the vector Qx.

Unitarily invariant matrix norm

Let $Q \in \mathbb{R}^{n\times n}$. Then Q is said to be orthogonal if $Q^\top Q = QQ^\top = I$. If Q is orthogonal then
$\|Qx\|_2 = \sqrt{\langle Qx, Qx\rangle} = \sqrt{\langle x, Q^\top Qx\rangle} = \sqrt{\langle x, x\rangle} = \|x\|_2.$

Let $Q \in \mathbb{C}^{n\times n}$. Then Q is said to be unitary if $Q^* Q = QQ^* = I$. If Q is unitary then $\|Qx\|_2 = \sqrt{\langle Qx, Qx\rangle} = \sqrt{\langle x, Q^* Qx\rangle} = \sqrt{\langle x, x\rangle} = \|x\|_2.$

It follows that $\|Q\|_2 = 1$ when Q is unitary and that error in x is NOT magnified in the vector Qx.

A matrix norm is said to be unitariliy invariant if $\|QAP\| = \|A\|$ holds for any unitary matrices Q and P.

Unitarily invariant matrix norm

Let $Q \in \mathbb{R}^{n \times n}$. Then Q is said to be orthogonal if $Q^\top Q = QQ^\top = I$. If Q is orthogonal then
$\|Qx\|_2 = \sqrt{\langle Qx,\, Qx\rangle} = \sqrt{\langle x,\, Q^\top Qx\rangle} = \sqrt{\langle x,\, x\rangle} = \|x\|_2.$

Let $Q \in \mathbb{C}^{n \times n}$. Then Q is said to be unitary if $Q^* Q = QQ^* = I$. If Q is unitary then $\|Qx\|_2 = \sqrt{\langle Qx,\, Qx\rangle} = \sqrt{\langle x,\, Q^* Qx\rangle} = \sqrt{\langle x,\, x\rangle} = \|x\|_2.$

It follows that $\|Q\|_2 = 1$ when Q is unitary and that error in x is NOT magnified in the vector Qx.

A matrix norm is said to be unitariliy invariant if $\|QAP\| = \|A\|$ holds for any unitary matrices Q and P. The spectral norm and the Frobenius norm are unitarily invariant norms.

Unitarily invariant matrix norm

Let $Q \in \mathbb{R}^{n\times n}$. Then Q is said to be orthogonal if $Q^\top Q = QQ^\top = I$. If Q is orthogonal then
$\|Qx\|_2 = \sqrt{\langle Qx,\, Qx\rangle} = \sqrt{\langle x,\, Q^\top Qx\rangle} = \sqrt{\langle x,\, x\rangle} = \|x\|_2.$

Let $Q \in \mathbb{C}^{n\times n}$. Then Q is said to be unitary if $Q^* Q = QQ^* = I$. If Q is unitary then $\|Qx\|_2 = \sqrt{\langle Qx,\, Qx\rangle} = \sqrt{\langle x,\, Q^* Qx\rangle} = \sqrt{\langle x,\, x\rangle} = \|x\|_2.$

It follows that $\|Q\|_2 = 1$ when Q is unitary and that error in x is NOT magnified in the vector Qx.

A matrix norm is said to be unitariliy invariant if $\|QAP\| = \|A\|$ holds for any unitary matrices Q and P. The spectral norm and the Frobenius norm are unitarily invariant norms. Indeed,

$$
\begin{aligned}
\|Qx\|_2 &= \|x\|_2 \Rightarrow \|QAP\|_2 = \|A\|_2. \\
\|QAP\|_F &= \sqrt{\mathrm{Tr}((QAP)^* QAP)} = \sqrt{\mathrm{Tr}(P^* A^* AP)} \\
&= \sqrt{\mathrm{Tr}(A^* A)} = \|A\|_F.
\end{aligned}
$$