Vector norm

A **vector norm** is a function $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ satisfying

1. $\|x\| \geq 0$ with equality iff $x = 0$,
2. $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{C}$, $x \in \mathbb{C}^n$,
3. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathbb{C}^n$.

Vector norm

A **vector norm** is a function $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ satisfying

1. $\|x\| \geq 0$ with equality iff $x = 0$,
2. $\|\alpha x\| = |\alpha| \|x\|$ for all $\alpha \in \mathbb{C}$, $x \in \mathbb{C}^n$,
3. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathbb{C}^n$.

Examples:

$$
\begin{aligned}
\|x\|_1 &= |x_1| + \cdots + |x_n|, \\
\|x\|_2 &= (|x_1|^2 + \cdots + |x_n|^2)^{1/2} = (x^* x)^{1/2}, \\
\|x\|_\infty &= \max_{1 \leq i \leq n} |x_i|, \\
\|x\|_p &= \left(|x_1|^p + \cdots + |x_n|^p\right)^{1/p}, \quad p \geq 1, \quad (p\text{-norm}).
\end{aligned}
$$

Vector norm

A **vector norm** is a function $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ satisfying

1. $\|x\| \geq 0$ with equality iff $x = 0$,
2. $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{C}$, $x \in \mathbb{C}^n$,
3. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathbb{C}^n$.

Examples:

$$\begin{aligned}
\|x\|_1 &= |x_1| + \cdots + |x_n|, \\
\|x\|_2 &= (|x_1|^2 + \cdots + |x_n|^2)^{1/2} = (x^* x)^{1/2}, \\
\|x\|_\infty &= \max_{1 \leq i \leq n} |x_i|, \\
\|x\|_p &= \left(|x_1|^p + \cdots + |x_n|^p\right)^{1/p}, \quad p \geq 1, \quad (p\text{-norm}).
\end{aligned}$$

Hölder inequality: $|x^* y| \leq \|x\|_p \|y\|_q$, $\frac{1}{p} + \frac{1}{q} = 1$.

Cauchy–Schwarz inequality: $|x^* y| \leq \|x\|_2 \|y\|_2$.

Vector norms (cont.)

All norms on $\mathbb{C}^n$ are **equivalent**: if $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are norms on $\mathbb{C}^n$ then there exist $\nu_1 > 0$ and $\nu_2 > 0$ such that

$$\nu_1 \|x\|_\alpha \leq \|x\|_\beta \leq \nu_2 \|x\|_\alpha, \quad \forall x \in \mathbb{C}^n.$$

Vector norms (cont.)

All norms on $\mathbb{C}^n$ are **equivalent**: if $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are norms on $\mathbb{C}^n$ then there exist $\nu_1 > 0$ and $\nu_2 > 0$ such that

$$\nu_1 \|x\|_\alpha \leq \|x\|_\beta \leq \nu_2 \|x\|_\alpha, \quad \forall x \in \mathbb{C}^n.$$

For example, for all $x \in \mathbb{C}^n$

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2 \leq n\|x\|_\infty.$$

Vector norms (cont.)

All norms on $\mathbb{C}^n$ are **equivalent**: if $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are norms on $\mathbb{C}^n$ then there exist $\nu_1 > 0$ and $\nu_2 > 0$ such that

$$\nu_1\|x\|_\alpha \leq \|x\|_\beta \leq \nu_2\|x\|_\alpha, \quad \forall x \in \mathbb{C}^n.$$

For example, for all $x \in \mathbb{C}^n$

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2 \leq n\|x\|_\infty.$$

The sequence $\{v^{(k)}\}$ of vectors in $\mathbb{C}^n$ **converges** to $v \in \mathbb{C}^n$ with respect to $\|\cdot\|$ iff $\lim_{k\to\infty}\|v^{(k)} - v\| = 0$.

Matrix norm

A matrix norm is a function $\|\cdot\| : \mathbb{C}^{m\times n} \to \mathbb{R}$ satisfying obvious analogues of the 3 vector norm properties.

- **Frobenius norm**:

$$\|A\|_F := \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2} = \big(\mathrm{trace}(A^*A)\big)^{1/2}.$$

Matrix norm

A matrix norm is a function $\|\cdot\| : \mathbb{C}^{m\times n} \to \mathbb{R}$ satisfying obvious analogues of the 3 vector norm properties.

- **Frobenius norm**:

$$\|A\|_F := \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2} = \big(\mathrm{trace}(A^*A)\big)^{1/2}.$$

- Matrix norms **subordinate** to vector norms:

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

- **matrix *p*-norm**: induced by the vector p-norm: $\|A\|_p := \max_{x\neq 0}\dfrac{\|Ax\|_p}{\|x\|_p}$.

Matrix p-norms

Theorem (Thm. 1)

Let $A \in \mathbb{C}^{m \times n}$. Then

$$\|A\|_1 = \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|, \quad \text{“max column sum”},$$

$$\|A\|_\infty = \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}| = \|A^T\|_1, \quad \text{“max row sum”},$$

$$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sqrt{\lambda_{\max}(A^* A)}, \quad \text{spectral norm},$$

where $\lambda_{\max}$ denotes the largest eigenvalue.

Consistent norms

A norm is **consistent** if it satisfies $\|AB\| \leq \|A\|\|B\|$.

The Frobenius norm and all subordinate norms are consistent.

Consistent norms

A norm is **consistent** if it satisfies $\|AB\| \leq \|A\|\|B\|$.

The Frobenius norm and all subordinate norms are consistent.

The “mixed" norm

$$\|A\|_{1,\infty} := \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_\infty} = \max_{i,j} |a_{ij}|$$

is a matrix norm but is not consistent. If $A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ then $\|AB\|_{1,\infty} = 2 > \|A\|_{1,\infty}\|B\|_{1,\infty} = 1 \times 1 = 1$.

Unitarily invariant norms

Let $A \in \mathbb{C}^{m \times n}$. For unitary $Q \in \mathbb{C}^{m \times m}$ and $Z \in \mathbb{C}^{n \times n}$,

$$\|QAZ\|_2 = \|A\|_2, \qquad \|QAZ\|_F = \|A\|_F.$$

$\Rightarrow$ $\|\cdot\|_2$ and $\|\cdot\|_F$ invariant under unitary transformations.

Unitarily invariant norms

Let $A \in \mathbb{C}^{m \times n}$. For unitary $Q \in \mathbb{C}^{m \times m}$ and $Z \in \mathbb{C}^{n \times n}$,

$$\|QAZ\|_2 = \|A\|_2, \qquad \|QAZ\|_F = \|A\|_F.$$

$\Rightarrow$ $\|\cdot\|_2$ and $\|\cdot\|_F$ invariant under unitary transformations.

If A is contaminated by errors E, $U(A+E)U^* = UAU^* + F$,

$$\|F\|_2 = \|UEU^*\|_2 = \|E\|_2.$$

$\Rightarrow$ multiplication by unitary matrices does not magnify errors.

Unitarily invariant norms

Let $A \in \mathbb{C}^{m\times n}$. For unitary $Q \in \mathbb{C}^{m\times m}$ and $Z \in \mathbb{C}^{n\times n}$,

$$\|QAZ\|_2 = \|A\|_2, \qquad \|QAZ\|_F = \|A\|_F.$$

$\Rightarrow$ $\|\cdot\|_2$ and $\|\cdot\|_F$ invariant under unitary transformations.

If A is contaminated by errors E, $U(A+E)U^* = UAU^* + F$,

$$\|F\|_2 = \|UEU^*\|_2 = \|E\|_2.$$

$\Rightarrow$ multiplication by unitary matrices does not magnify errors.

For a general, nonsingular similarity transformation

$$X(A+E)X^{-1} = XAX^{-1} + G,$$

$\|G\|_2 = \|XEX^{-1}\|_2 \le \kappa_2(X)\|E\|_2$, where $\kappa_2(X) = \|X\|_2\|X^{-1}\|_2$ is the **condition number** of X and can be arbitrarily large.

Semyon Aranovich Gershgorin (1901–1933)

Theorem (Thm. 4, Gershgorin (1931))

The eigenvalues of $A \in \mathbb{C}^{n\times n}$ lie in the union of the n discs in the complex plane

$$D_i = \Big\{ z \in \mathbb{C} : |z - a_{ii}| \le \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \Big\}, \qquad i = 1, \dots, n.$$

Example

Let $A = \begin{bmatrix} 8 & 1 & 0 \\ 1 & 10 & 1 \\ 0 & 1 & 12 \end{bmatrix}$.

What can be said about the location of A's eigenvalues?