CS 450: Numerical Anlaysis

Lecture 3
Chapter 2 – Linear Systems
Matrix Norms and Conditioning

Edgar Solomonik

Department of Computer Science
University of Illinois at Urbana-Champaign

Vector Norms

n-dim space

$n=3$

- **Properties of vector norms**

$\|x\| = 0$ if and only if $x = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

$\|\alpha x\| = |\alpha| \cdot \|x\|$

$\|x\| \geq 0$, $\|x+y\| \leq \|x\| + \|y\|$

- **A norm is uniquely defined by its unit sphere:**

$V \subset \mathbb{R}^n$, $x \in V$, $\|x\| = 1$

- **p-norms**

$\|x\|_1 = \sum_i |x_i|$

$\|x\|_2 = \sqrt{\sum_i x_i^2}$

$\|x\|_p = \left(\sum_i |x_i|^p\right)^{1/p}$

$\|x\|_\infty = \max_i (|x_i|)$

Inner-Product Spaces

- **Properties of inner-product spaces**: Inner products $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ must satisfy

e.g.

$\langle x, y \rangle = x^T y$

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$$
$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{0}$$
$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$$
$$\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$$
$$\langle \alpha \boldsymbol{x}, \boldsymbol{y} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle$$

- **Inner-product-based vector norms**

$\|x\|_2 = x^T x$

more generally, for any $\langle x, y \rangle \Rightarrow x^T A y \Rightarrow$ new norm

$x^T A x$

$\langle x, x \rangle$

Cauchy-Schwartz inequality : $\Rightarrow |x^T y| \leq \|x\|_2 \|y\|_2$

$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \langle y, y \rangle} \Rightarrow \|x\|_2 \|y\|_2$

Matrix Norms

- **Properties of matrix norms:**

$A \in \mathbb{R}^{m \times n}$

sometimes have submultiplicativity

$\|A\| \|B\| \geq \|AB\|$

$$\|\boldsymbol{A}\| \geq 0$$
$$\|\boldsymbol{A}\| = 0 \quad \Leftrightarrow \quad \boldsymbol{A} = \boldsymbol{0}$$
$$\|\alpha \boldsymbol{A}\| = |\alpha| \cdot \|\boldsymbol{A}\|$$
$$\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\| \quad (\textit{triangle inequality})$$

- **Frobenius norm:** $\|A\|_F = \sqrt{\sum_i \sum_j a_{ij}^2} = \|\text{vec}(A)\|_2$

stack col's of A into vector

- **Operator/induced/subordinate matrix norms:**

$\hookrightarrow$ from vector p-norm

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

Induced Matrix Norms

- **Interpreting induced matrix norms:**

maximum amplification of any vector

$$y = Ax \Rightarrow \|y\| \le \|A\| \cdot \|x\|$$

$$A^{-1}y = x \Rightarrow \|x\| \le \|A^{-1}\| \cdot \|y\|$$

$$\|y\| \ge \|x\| / \|A^{-1}\|$$

- **General induced matrix norms:**

$$\|A\|_{pq} = \max_{\|x\|_p = 1} \|Ax\|_q$$

Matrix Condition Number

$\kappa(A) = ||A||_2 ||A^{-1}||_2$

- **Definition**: $\kappa(\boldsymbol{A}) = ||\boldsymbol{A}|| \cdot ||\boldsymbol{A}^{-1}||$ is the ratio between the shortest/longest distances from the unit-ball center to any point on the surface.

- **Intuitive derivation**:

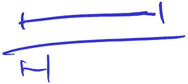
$||A||$

$$\kappa(\boldsymbol{A}) = \max_{\text{inputs}} \lim_{||\delta x|| \to 0} \max_{\text{perturbations in input}} \left| \frac{\text{relative perturbation in output}}{\text{relative perturbation in input}} \right|$$

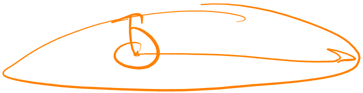
$1/||A^{-1}||$

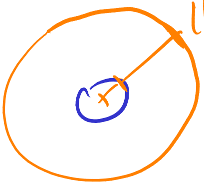
since a matrix is a linear operator, we can decouple its action on the input $\boldsymbol{x}$ and the perturbation $\boldsymbol{\delta x}$ since $\boldsymbol{A}(\boldsymbol{x} + \boldsymbol{\delta x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}\boldsymbol{\delta x}$, so

$$\kappa(\boldsymbol{A}) = \left| \frac{\overbrace{\max_{\text{perturbations in input}} \text{relative perturbation growth}}^{||\boldsymbol{A}||}}{\underbrace{\max_{\text{inputs}} \text{relative input reduction}}_{1/||\boldsymbol{A}^{-1}||}} \right|$$



$= \kappa(A)$





$\|A\| > 1$

$\kappa(A) = 1$