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

January 26, 2018

Vector Norms

- **Properties of vector norms**

$$
\begin{aligned}
||\boldsymbol{x}|| &\geq 0 \\
||\boldsymbol{x}|| &= 0 \quad \Leftrightarrow \quad \boldsymbol{x} = \mathbf{0} \\
||\alpha \boldsymbol{x}|| &= |\alpha| \cdot ||\boldsymbol{x}|| \\
||\boldsymbol{x} + \boldsymbol{y}|| &\leq ||\boldsymbol{x}|| + ||\boldsymbol{y}||
\end{aligned}
$$

(triangle inequality) implies continuity

- p**-norms**

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

Inner-Product Spaces

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

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$$

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{0}$$

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$$

$$\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$$

$$\langle \alpha \boldsymbol{x}, \boldsymbol{y} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle$$

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

$$\sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$$

is an norm, due to Cauchy-Schwartz inequality

$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle \cdot \langle \boldsymbol{y}, \boldsymbol{y} \rangle}$$

Inner-products can be expressed as $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{y}$ *where* $\boldsymbol{A}$ *is symmetric positive definite, yielding norms* $||\boldsymbol{x}||_{\boldsymbol{A}} = \sqrt{\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}}$

Matrix Norms

- **Properties of matrix norms**:

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

- **Frobenius norm**:

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

- **Operator/induced/subordinate matrix norms**:
  For any vector norm $|| \cdot ||$, *the induced matrix norm is*

$$
||\boldsymbol{A}|| = \max_{\boldsymbol{x} \neq \boldsymbol{0}} ||\boldsymbol{A}\boldsymbol{x}|| / ||\boldsymbol{x}|| = \max_{||\boldsymbol{x}||=1} ||\boldsymbol{A}\boldsymbol{x}||
$$

Induced Matrix Norms

- **General induced matrix norms**:

$$||\boldsymbol{A}||_{mp} = \max_{||\boldsymbol{x}||_p=1} ||\boldsymbol{A}\boldsymbol{x}||_m$$

typically $m = p$ *so we write* $||\boldsymbol{A}||_p$ *and almost always we have* $p \in \{1, 2, \infty\}$. *(Computing the matrix norm for certain choices of* $m \neq p$ *is NP-complete.)*

- **Interpreting induced matrix norms**:

$$||\boldsymbol{A}||_p = \max_{||\boldsymbol{x}||_p=1} ||\boldsymbol{A}\boldsymbol{x}||_p$$

is the maximum possible p*-norm amplification due to application of* $\boldsymbol{A}$

$$1/||\boldsymbol{A}^{-1}||_p = \min_{||\boldsymbol{x}||_p=1} ||\boldsymbol{A}\boldsymbol{x}||_p$$

is the maximum possible p*-norm reduction due to application of* $\boldsymbol{A}$

Matrix Condition Number

- **Definition**: $\kappa(\boldsymbol{A}) = ||\boldsymbol{A}|| \cdot ||\boldsymbol{A}^{-1}||$
- **Intuitive derivation**:

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

since a matrix is a linear operator, we can generally decouple its action on the input and the perturbation, so

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