CS 450: Numerical Anlaysis[1]

Linear Systems

$$\underline{A}\,\underset{?}{x} = \underline{b}$$

University of Illinois at Urbana-Champaign

---

[1] *These slides have been drafted by Edgar Solomonik as lecture templates and supplementary material for the book "Scientific Computing: An Introductory Survey" by Michael T. Heath (slides).*

Vector Norms

$x \in \mathbb{R}^n$

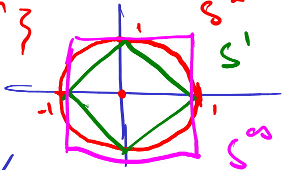
- **Properties of vector norms**

$\|x\| \geq 0$, $\|x\| = 0$ if $x = \begin{bmatrix} \vdots \\ 0 \\ \vdots \end{bmatrix}$

$\|\alpha x\| = |\alpha| \|x\|$, $\|x+y\| \leq \|x\| + \|y\|$

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

$S^p = \{x : \|x\|_p = 1, x \in \mathbb{R}^n\}$



$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$

$\|x\|_1 = \sum_{i=1}^n |x_i|$

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

- ***p*-norms**

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

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**

$\|x\|_2 = \sqrt{\langle x, x \rangle}$

$= \sum_i x_i y_i$

commutativity

distributivity

$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \langle y, y \rangle}$, with $\langle x, x \rangle = \|x\|_2^2$

$|\langle x, y \rangle| \leq \|x\|_2 \|y\|_2$

Matrix Norms

- **Properties of matrix norms:**

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

$$\mathrm{vec}\left(\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}\right) \to \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

- **Frobenius norm:** $\|A\|_F = \sqrt{\sum_{i=1} \sum_{j=1} a_{ij}^2} = \|\mathrm{vec}(A)\|_2$

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

$$A \to f(x) = Ax$$

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

Induced Matrix Norms

- **Interpreting induced matrix norms:**

$\|A\|_p \rightarrow$ maximum ampl. factor for the norm of Ax relative to x over all x

$$\min_{\substack{x \in \mathbb{R}^n \\ \|x\|_p = 1}} \|Ax\|_p = \frac{1}{\|A^{-1}\|_p} \quad \Big| \quad y = Ax, \quad A^{-1} y = x$$

- **General induced matrix norms:**

Matrix Condition Number

***Demo:** Conditioning of 2x2 Matrices*
***Demo:** Condition number visualized*

- **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**:

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

since a matrix is a linear operator, we can 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

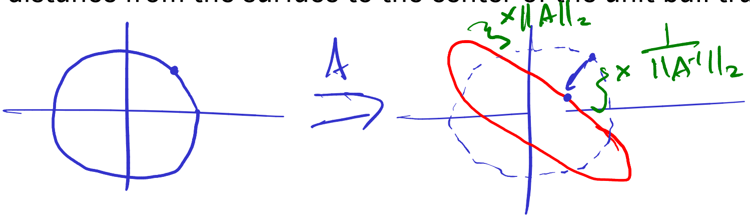
$$||A\delta x|| \le ||A||_2 ||\delta x||$$

$$||Ax|| \ge \frac{1}{||A^{-1}||} ||x||$$

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

Matrix Conditioning

- The matrix condition number $\kappa(\boldsymbol{A})$ is the ratio between the max and min distance from the surface to the center of the unit ball transformed by $\kappa(\boldsymbol{A})$:



$$A^{-1}y = x$$

- The matrix condition number bounds the worst-case amplification of error in a matrix-vector product:

$$y + \delta y = A(x + \delta x) = Ax + A\delta x$$

$$\delta y = A\delta x$$

$$\|\delta y\|_2 \le \|A\|_2 \|\delta x\|_2$$

$$\|y\|_2 \ge \frac{1}{\|A^{-1}\|_2}\|x\|$$

$$\frac{\|\delta y\|}{\|y\|} \overset{?}{\le} \frac{\|\delta x\|}{\|x\|} \quad \|A\|_2\|A^{-1}\|_2$$

Norms and Conditioning of Orthogonal Matrices

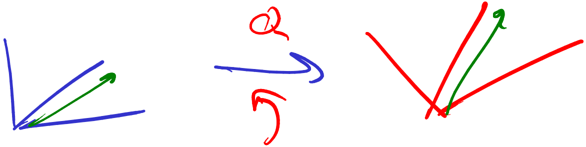
- **Orthogonal matrices:**

$Q^T = Q^{-1}$ $\quad Q^TQ = I$ $\quad QQ^T = I$

$Q \in \mathbb{R}^{n\times n}$

- **Norm and condition number of orthogonal matrices:**

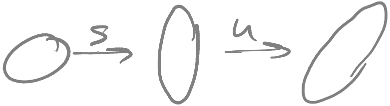
$\|Q\|_2 = 1$ $\quad \|Qx\|_2 = \|x\|_2$



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

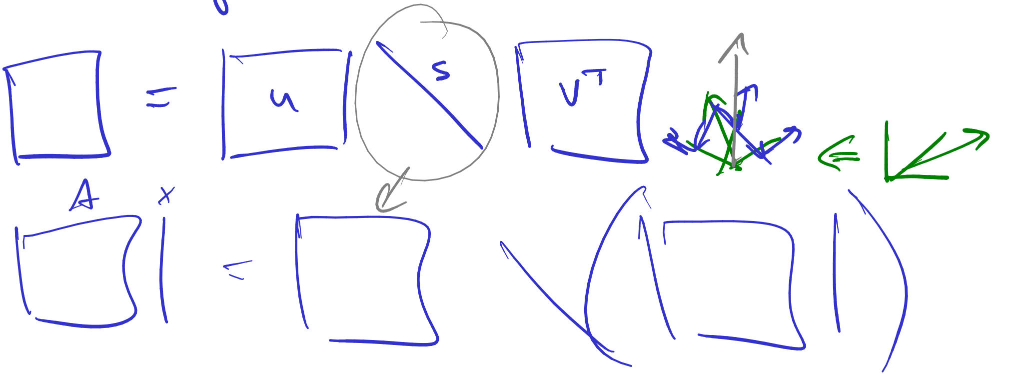
$Q^{-1} \to Q^T$

Singular Value Decomposition



- **The singular value decomposition (SVD):**

$A = U S V^T$ — S: diagonal and nonnegative (positive if A is full rank); U, V: orthogonal



Norms and Conditioning via SVD

- **Norm and condition number in terms of singular values:**

$$S = \begin{bmatrix} \sigma_{max} & & \\ & \ddots & \\ & & \sigma_{min} \end{bmatrix}$$

singular values

$$S = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$$

$$\max_{\|x\|_2=1} \|Sx\|_2 = \max_{\|x\|_2=1} \sqrt{\sum_i \sigma_i^2 x_i^2}$$

$$\|S\|_2 = \sigma_{max} \qquad \|S^{-1}\|_2 = \frac{1}{\sigma_{min}}$$

$$\Rightarrow \|A\| \le \|U\|\,\|S\|\,\|V^T\|$$

$$A = USV^T$$

$$U^T A V = S$$

$$\|S\| \le \|U^T\|\,\|A\|\,\|V\|$$

$$\|A\|\,\|A^{-1}\| = \|S\|\,\|S^{-1}\| = \sigma_{max}/\sigma_{min}$$

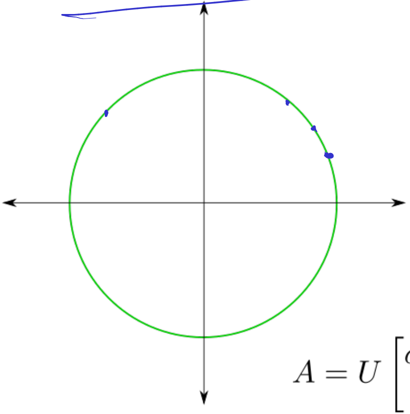
Visualization of Matrix Conditioning

$(A^TA)^T = A^TA$

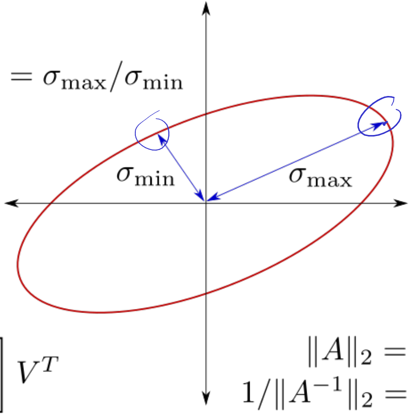
$\frac{x^TA^T(Ax)}{x^Tx} \geq 0$

$\frac{y^Ty}{x^Tx} \geq 0$

$\{x : x \in \mathbb{R}^2, \|x\|_2 = 1\} \xrightarrow{A} \{Ax : x \in \mathbb{R}^2, \|x\|_2 = 1\}$



$\kappa(A) = \sigma_{\max}/\sigma_{\min}$



$$A = U \begin{bmatrix} \sigma_{\max} & \\ & \ddots & \\ & & \sigma_{\min} \end{bmatrix} V^T$$

$$\|A\|_2 = \sigma_{\max}$$

$$1/\|A^{-1}\|_2 = \sigma_{\min}$$

$$A^{-1} = V \begin{bmatrix} \frac{1}{\sigma_{max}} & \\ & \ddots & \\ & & \frac{1}{\sigma_{min}} \end{bmatrix} U^T$$

Conditioning of Linear Systems

- **Lets now return to formally deriving the conditioning of solving $Ax = b$:**

Conditioning of Linear Systems II

- **Consider perturbations to the input coefficients $\hat{A} = A + \delta A$:**

Solving Basic Linear Systems

- Solve $\boldsymbol{D}\boldsymbol{x} = \boldsymbol{b}$ if $\boldsymbol{D}$ is diagonal
- Solve $\boldsymbol{Q}\boldsymbol{x} = \boldsymbol{b}$ if $\boldsymbol{Q}$ is orthogonal
- Given SVD $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$, solve $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

Solving Triangular Systems

- $Lx = b$ **if** L **is lower-triangular is solved by forward substitution:**

$$\begin{aligned} l_{11}x_1 &= b_1 & & x_1 = \\ l_{21}x_1 + l_{22}x_2 &= b_2 & \Rightarrow \quad & x_2 = \\ l_{31}x_1 + l_{32}x_2 + l_{33}x_3 &= b_3 & & x_3 = \\ &\vdots & & \vdots \end{aligned}$$

- **Algorithm can also be formulated recursively by blocks:**

Solving Triangular Systems

- **Existence of solution to $Lx = b$:**

- **Uniqueness of solution:**

- **Computational complexity of forward/backward substitution:**

Properties of Triangular Matrices

- **$Z = XY$ is lower triangular is X and Y are both lower triangular:**

- **L^{-1} is lower triangular if it exists:**

LU Factorization

- **An *LU factorization* consists of a unit-diagonal lower-triangular *factor* L and upper-triangular factor U such that $A = LU$:**

- **Given an LU factorization of A, we can solve the linear system $Ax = b$:**

Gaussian Elimination Algorithm

Demo: *LU factorization*

- **Algorithm for factorization is derived from equations given by $A = LU$:**

- **The computational complexity of LU is $O(n^3)$:**

Existence of LU Factorization

- **The LU factorization may not exist:** Consider matrix $\begin{bmatrix} 3 & 2 \\ 6 & 4 \\ 0 & 3 \end{bmatrix}$.

- **Permutation of rows enables us to transform the matrix so the LU factorization does exist:**

Gaussian Elimination with Partial Pivoting

- ***Partial pivoting*** **permutes rows to make divisor u_{ii} is maximal at each step:**

- **A row permutation corresponds to an application of a *row permutation matrix* $\boldsymbol{P}_{jk} = \boldsymbol{I} - (\boldsymbol{e}_j - \boldsymbol{e}_k)(\boldsymbol{e}_j - \boldsymbol{e}_k)^T$:**

Partial Pivoting Example

- Lets consider again the matrix $\boldsymbol{A} = \begin{bmatrix} 3 & 2 \\ 6 & 4 \\ 0 & 3 \end{bmatrix}$.

Complete Pivoting

- ***Complete pivoting*** **permutes rows and columns to make divisor u_{ii} is maximal at each step:**

- **Complete pivoting is noticeably more expensive than partial pivoting:**

Round-off Error in LU

- **Lets consider factorization of $\begin{bmatrix} \epsilon & 1 \\ 1 & 1 \end{bmatrix}$ where $\epsilon < \epsilon_{\text{mach}}$:**

- **Permuting the rows of A in partial pivoting gives $PA = \begin{bmatrix} 1 & 1 \\ \epsilon & 1 \end{bmatrix}$**

Error Analysis of LU

- **The main source of round-off error in LU is in the computation of the Schur complement:**

- **When computed in floating point, absolute backward error δA in LU (so $\hat{\boldsymbol{L}}\hat{\boldsymbol{U}} = \boldsymbol{A} + \boldsymbol{\delta A}$) is $|\delta a_{ij}| \leq \epsilon_{\text{mach}}(|\hat{\boldsymbol{L}}| \cdot |\hat{\boldsymbol{U}}|)_{ij}$**

Helpful Matrix Properties

- **Matrix is *diagonally dominant*, so $\sum_{i \neq j} |a_{ij}| \leq |a_{ii}|$:**

- **Matrix is *symmetric positive definite (SPD)*, so $\forall_{\boldsymbol{x} \neq 0}, \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} > 0$:**

- **Matrix is symmetric but indefinite:**

- **Matrix is *banded*, $a_{ij} = 0$ if $|i - j| > b$:**

Solving Many Linear Systems

- **Suppose we have computed $A = LU$ and want to solve $AX = B$ where B is $n \times k$ with $k < n$:**

- **Suppose we have computed $A = LU$ and now want to solve a perturbed system $(\boldsymbol{A} - \boldsymbol{u}\boldsymbol{v}^T)\boldsymbol{x} = \boldsymbol{b}$:**
  Can use the *Sherman-Morrison-Woodbury* formula

$$(\boldsymbol{A} - \boldsymbol{u}\boldsymbol{v}^T)^{-1} = \boldsymbol{A}^{-1} + \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^T\boldsymbol{A}^{-1}}{1 - \boldsymbol{v}^T\boldsymbol{A}^{-1}\boldsymbol{u}}$$