CS450: Numerical Analysis

Howard Hao Yang

Assistant Professor, ZJU-UIUC Institute

08/10/2023

Linear Systems

- We will concentrate on the following problems in the next few weeks
  - Linear system problem: $\boldsymbol{Ax} = \boldsymbol{b}$, find $\boldsymbol{x}$
  - Eigen vector problem: $\boldsymbol{Ax} = \lambda \boldsymbol{x}$, find $\boldsymbol{x}$ and λ
  - Singular value decomposition: $\boldsymbol{Av} = \sigma \boldsymbol{u}$, find $\boldsymbol{v}$, $\boldsymbol{u}$ and σ
  - Factor the matrix as $\boldsymbol{A} = \boldsymbol{CR}$
- These are central problems in linear algebra as well as data science
- Let's begin with a quick review of linear algebra
  - Main focus: column space and ranks

Matrix-Vector Multiplication

- Example: Multiply $\boldsymbol{A}$ times $\boldsymbol{x}$

$$\textbf{By rows} \quad \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ 2x_1 + 4x_2 \\ 3x_1 + 7x_2 \end{bmatrix} = \begin{array}{c} \text{inner products} \\ \text{of the rows} \\ \text{with } \boldsymbol{x} = (x_1, x_2) \end{array}$$

$$\textbf{By columns} \quad \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix} = \begin{array}{c} \text{combination} \\ \text{of the columns} \\ \boldsymbol{a}_1 \text{ and } \boldsymbol{a}_2 \end{array}$$

  - Perspective #1: produce three inner/dot products by each row, useful for computing, but not for understanding
  - Perspective #2: linear combination of $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$
  - In essence, $\boldsymbol{Ax}$ is a linear combination of columns of $\boldsymbol{A}$

Column Space

- Definition (column space): The combination of the columns fill out the column space of $\boldsymbol{A}$. In other words, the possible outcome of $\boldsymbol{Ax}$ when $\boldsymbol{x}$ goes through all possible values: $\{\boldsymbol{Ax} | \boldsymbol{x} \in \boldsymbol{R}^d\}$.
  - Example: The possible subspaces of $\boldsymbol{R}^3$

    The **zero vector** $(0,0,0)$ by itself
    A **line** of all vectors $x_1\boldsymbol{a}_1$
    A **plane** of all vectors $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2$
    The **whole** $\mathbf{R}^3$ with all vectors $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3$

  - Exercise: What are the column spaces of $A_2 = \begin{bmatrix} 2 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 10 \end{bmatrix}$ and $A_3 = \begin{bmatrix} 2 & 3 & 1 \\ 2 & 4 & 1 \\ 3 & 7 & 1 \end{bmatrix}$?

Independent Columns and Rank

- Goal: Given matrix $\boldsymbol{A}$, construct a matrix $\boldsymbol{C}$ where the columns come directly from $\boldsymbol{A}$, but not to include any column that is a combination of the previous ones, following is the approach:
  - If column 1 of $\boldsymbol{A}$ is not all zero, put it into the matrix $\boldsymbol{C}$;
  - If column 2 of $\boldsymbol{A}$ is not a multiple of column 1, put it into $\boldsymbol{C}$;
  - If column 3 of $\boldsymbol{A}$ is not a combination of columns 1 and 2, put it into $\boldsymbol{C}$. Continue.
  - At the end, $\boldsymbol{C}$ will have r columns, where $r \le n$.
  - They will form a basis for the column space of $\boldsymbol{A}$.
  - All vectors in the space are combinations of the basis vectors.

Independent Columns and Rank (cont'd)

- Examples:
  - If $A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix}$ then $C = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}$ $\quad n = 3$ columns in A, $\boldsymbol{r} = \mathbf{2}$ columns in C
  - If $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$ then $C = A$. $\quad n = 3$ columns in A, $\boldsymbol{r} = \mathbf{3}$ columns in C
  - If $A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 2 & 5 \\ 1 & 2 & 5 \end{bmatrix}$ then $C = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ $\quad n = 3$ columns in A, $\boldsymbol{r} = \mathbf{1}$ column in C

Independent Columns and Rank (cont'd)

- The number of columns r in matrix $\boldsymbol{C}$ is the "rank" of matrix $\boldsymbol{A}$.
- The rank counts independent columns, i.e., the rank of a matrix is the dimension of its column space.
- We can construct different basis, but always the same number of vectors.
- The matrix $\boldsymbol{C}$ connects matrix $\boldsymbol{A}$ with another matrix, as $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{R}$; a factorization operation:

$$\boxed{\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix} = \boldsymbol{C}\boldsymbol{R}}$$

Matrix-Matrix Multiplication

- Example: Multiply $\boldsymbol{A}$ times $\boldsymbol{B}$, as $\boldsymbol{AB} = \boldsymbol{C}$
  - Perspective #1: inner product approach, use row of $\boldsymbol{A}$ multiply column of $\boldsymbol{B}$, facilitates computations

row 2 of A
column 3 of B
give c_{23} in C

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ a_{21} & a_{22} & a_{23} \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & \cdot & b_{13} \\ \cdot & \cdot & b_{23} \\ \cdot & \cdot & b_{33} \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & c_{23} \\ \cdot & \cdot & \cdot \end{bmatrix}$$

  - Perspective #2: use column of $\boldsymbol{A}$ multiply with row of $\boldsymbol{B}$, facilitates understanding

Matrix-Matrix Multiplication (cont'd)

- Outer product: one column matrix $\boldsymbol{u}$ times $\boldsymbol{B}$, as $\boldsymbol{AB} = \boldsymbol{C}$
  - A column matrix $\boldsymbol{u}$ times a row matrix $\boldsymbol{v}^{\boldsymbol{T}}$ produces a matrix

"Outer product" $$uv^{\mathrm{T}} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 8 & 12 \\ 6 & 8 & 12 \\ 3 & 4 & 6 \end{bmatrix} =$$ **"rank one matrix"**

  - The column space of $\boldsymbol{uv}^{\boldsymbol{T}}$ is one-dimensional: the line in the same direction of $\boldsymbol{u}$
  - All nonzero matrices $\boldsymbol{uv}^{\boldsymbol{T}}$ have rank one – they are the perfect building blocks for every matrix

Matrix-Matrix Multiplication (cont'd)

- Write the product $\boldsymbol{AB}$ as a sum of rank-one matrices
  - Column-row multiplication of matrices

$$AB = \begin{bmatrix} | & & | \\ \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} — & \boldsymbol{b}_1^* & — \\ & \vdots & \\ — & \boldsymbol{b}_n^* & — \end{bmatrix} = \underbrace{\boldsymbol{a}_1 \boldsymbol{b}_1^* + \boldsymbol{a}_2 \boldsymbol{b}_2^* + \cdots + \boldsymbol{a}_n \boldsymbol{b}_n^*}_{\textbf{sum of rank 1 matrices}}.$$

  - Example

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{3} \end{bmatrix} \begin{bmatrix} \mathbf{2} & \mathbf{4} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{5} \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 12 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 17 \end{bmatrix}$$

  - Why writing it as this form is important?

Ranks of AB and $A + B$

- Let's think about the following relationships
  - When we multiply matrices, can we increase the rank?
  - In other words, would it hold?: $\text{rank}(\boldsymbol{AB}) > \text{rank}(\boldsymbol{A})$
  - When we sum up matrices, can we increase the rank?
  - In other words, would it hold?: $\text{rank}(\boldsymbol{A} + \boldsymbol{B}) > \text{rank}(\boldsymbol{A})$

Ranks of AB and $A + B$ (cont'd)

- Important inequalities for ranks
  - When we multiply matrices, we cannot increase the rank:
    - $\text{rank}(\boldsymbol{AB}) \leq \text{rank}(\boldsymbol{A})$
    - $\text{rank}(\boldsymbol{AB}) \leq \text{rank}(\boldsymbol{B})$
  - Rank of summations
    - $\text{rank}(\boldsymbol{A} + \boldsymbol{B}) \leq \text{rank}(\boldsymbol{A}) + \text{rank}(\boldsymbol{B})$
  - Given matrix $\boldsymbol{A}$, the rank of $\boldsymbol{A}^T\boldsymbol{A}$ satisfies
    - $\text{rank}\left(\boldsymbol{A}^T\boldsymbol{A}\right) = \text{rank}\left(\boldsymbol{A}\boldsymbol{A}^T\right) = \text{rank}(\boldsymbol{A})$

Four Fundamental Subspaces

- The following subspaces are essential in characterizing $\boldsymbol{A}$
  - The column space $\mathbf{C}(\boldsymbol{A})$ contains all combinations of the columns of $\boldsymbol{A}$
  - The row space $\mathbf{C}(\boldsymbol{A}^T)$ contains all combinations of the columns of $\boldsymbol{A}^T$
  - The nullspace $\mathbf{N}(\boldsymbol{A})$ contains all solutions $\boldsymbol{x}$ to $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$
  - The left nullspace $\mathbf{N}(\boldsymbol{A}^T)$ contains all solutions $\boldsymbol{y}$ to $\boldsymbol{A}^T\boldsymbol{y} = \mathbf{0}$

  - Example: The null space

$$Bx = \begin{bmatrix} 1 & -2 & -2 \\ 3 & -6 & -6 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \textbf{ has solutions } x_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \textbf{ and } x_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

Four Fundamental Subspaces (cont'd)

- Exercise
  - Suppose matrix $\boldsymbol{A}$ is a 3-by-3 matrix of all ones, find two independent vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ that solves $\boldsymbol{Ax} = \mathbf{0}$ and $\boldsymbol{Ay} = \mathbf{0}$ (note: $\boldsymbol{x}$ and $\boldsymbol{y}$ shall be non-trivial, i.e., $\boldsymbol{x} \neq \boldsymbol{y} \neq \mathbf{0}$).
  - Why don't I ask for a third independent vector that solves $\boldsymbol{Az} = \mathbf{0}$? What does this imply?

Vector Norms

- What is a (vector) norm?
- A metric to measure the "length" of a vector, or "distance" between two vectors
  - An operator that $\|\cdot\|: R^d \rightarrow R_+$, and satisfies
    - $\|\boldsymbol{x}\| \geq 0$ and $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = \mathbf{0}$
    - $\|a \cdot \boldsymbol{x}\| = |a| \cdot \|\boldsymbol{x}\|$ for any $a \in R$
    - $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

Vector Norms (cont'd)

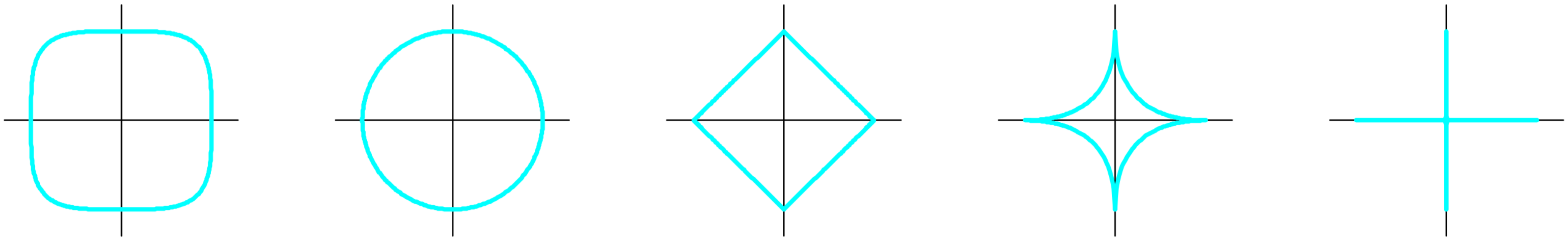
- Typical norms for vectors
  - The L-p norms
    - L-p norm: $\|\boldsymbol{\beta}\|_p = \left(\sum_{j=1}^{d} \beta_j^p\right)^{1/p}$, $p \geq 1$
    - $p = 1$, $\|\boldsymbol{\beta}\|_1 = \sum_j |\beta_j|$
    - $p = 2$, $\|\boldsymbol{\beta}\|_2 = \sqrt{\sum_j |\beta_j|^2}$
    - $p = \infty$, $\|\boldsymbol{\beta}\|_\infty = \max_{\mathrm{j}} |\beta_j|$
    - The L-0 norm: Counts the number of non-zero entries, e.g., if $\boldsymbol{\beta} = (10, 0, 2, 0.01, 0, 1)^T$, then $\|\boldsymbol{\beta}\|_0 = 4$

Vector Norms (cont'd)

- Exercise
  - For vector $\boldsymbol{x} = [-1.6, 1.2]^T$, calculate the L-1 norm, L-2 norm, and L-∞
  - In general, for any vector $\boldsymbol{x}$, does it hold that $\|\boldsymbol{x}\|_1 \geq \|\boldsymbol{x}\|_2 \geq \|\boldsymbol{x}\|_\infty$?

Side-Track: Unit Ball

- Unit ball under different norms
  - L-p norm: $\|\boldsymbol{\beta}\|_p = \left(\sum_{j=1}^{d} \beta_j^p\right)^{1/p}$
  - Unit ball under L-p norm: $\{\boldsymbol{\beta} \in R^d : \|\boldsymbol{\beta}\|_p = 1\}$



$p = 4$ $\quad p = 2$ $\quad p = 1$ $\quad p = 0.5$ $\quad p = 0.1$

Matrix Norms

- Given matrix $\boldsymbol{A}$
  - We want an operator that satisfies
    - $\|\boldsymbol{A}\| \geq 0$ and $\|\boldsymbol{A}\| = 0$ if and only if $\boldsymbol{A} = \boldsymbol{0}$
    - $\|a \cdot \boldsymbol{A}\| = |a| \cdot \|\boldsymbol{A}\|$ for any $a \in R$
    - $\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$
  - How to achieve this?
    - The Frobenius norm

Matrix Norms (cont'd)

- Given a vector norm $||\cdot||$
  - The matrix norm induced by this vector norm is given as

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

    - Different vector norms can induce different matrix norms for the same matrix
  - Examples
    - When the norm is L-1 norm, $||\boldsymbol{A}||_{\boldsymbol{1}} = \max_{j} \sum_{i=1}^{m} |a_{ij}|$: the maximum absolute column sum of the matrix
    - When the norm is L-∞ norm, $||\boldsymbol{A}||_{\infty} = \max_{i} \sum_{j=1}^{n} |a_{ij}|$: the maximum absolute row sum of the matrix

Matrix Norms (cont'd)

- Exercise
  - What are the matrix norms induced by L-1 norm and L-∞ norm for the following matrix?

$$\boldsymbol{A} = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 3 & -1 & 4 \end{bmatrix}$$

Matrix Norms (cont'd)

- When the norm is the Euclidean norm $\|\cdot\|_2$
  - The matrix norm induced by this vector norm is given as

$$\|\boldsymbol{A}\|_{\mathbf{2}} = \max_{\boldsymbol{x}\neq\mathbf{0}} \frac{\|\boldsymbol{Ax}\|_2}{\|\boldsymbol{x}\|_2} = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{Ax}\|_{\mathbf{2}}$$

    - This is called the spectral norm of matrix $\boldsymbol{A}$

Matrix Norm

- What does matrix multiplication do to a linear (sub)space?
  - Rotation & stretching
- What does matrix norm mean?
  - How much a particular linear operator A stretch a space