Linear Algebra?

Lecture - Prof Gilbert Strang - 18.06SC

Why Linear Algebra?

- Data will be mostly in form of Matrices and Vectors.
- Neural Network training can be seen as a series of Matrix Multiplication
- Why? to scale under techniques suitable for matrix based exchanges

Topics

- Vectors & matrices
- Dot product
- Matrix - vector (matrix matrix multiplication
- Rank, null space and column space
- Matrix Decomposition

Vectors

- Usually will represent a data point for example a house, an image, a document
- $\begin{bmatrix} 2.3 \\ 3.5 \end{bmatrix}$ May represent a person with age, weight 65 kg & height 185 cm
  If its said to be in $\mathbb{R}^n$
- Usually a vector with 2 coordinates (x-axis and y-axis) is said to be in $\mathbb{R}^2$
- The coordinates may not also be integer/whole
- We can visualize vectors upto 3 dimensions

$\|v_1\|_2 = \sqrt{x^2+y^2}$
$= \sqrt{2^2+3^2}$
$= \sqrt{13}$
$= 3.605$

- Can be added coordinate wise

Matrices:-

- Represent full data matrices

$$\begin{bmatrix} \text{data} \\ \text{points} \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \leftarrow \text{event}$$

- May represent weights of a neural network May represent a linear transformation.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Frobenius Norm

$\|A\|_F$

For $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

$$\|A\|_F = \sqrt{1^2+2^2+3^2+4^2}$$

Dot Product

$a = [1 \;\; 2 \;\; 3]^T$

$b = [2 \;\; 4 \;\; 6]^T$

$a \cdot b = a^T b$

$= [1 \;\; 2 \;\; 3] \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$

$= (1 \times 2) + 2 \cdot 4 + 3 \cdot 6$
$= 2 + 8 + 18$
$= 28$

(Cosine relation and scalar body)

Matrix Multiplication :-

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} (1\times5)+(2\times7) & (1\times6)+(2\times8) \\ (3\times5)+(4\times7) & (3\times6)+(4\times8) \end{bmatrix}$$

$$= \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

(Necessary condition for matrix multiplication)

In general for

$$\underset{m}{A} \; \underset{n}{B} \; \underset{p}{C} = \underset{m}{D} \; \underset{p}{E}$$

$$\underset{m}{A} \; \underset{n}{B} \; \underset{p}{C} \; \underset{q}{D}$$

Matrix Vector Multiplication

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$x_1 (a_{11}, a_{21}, a_{31}) + x_2 (a_{12}, a_{22}, a_{32}) + \dots x_n (a_{1n}, a_{2n}, a_{3n})$

- Notions of linear Independence, Rank & Column Space

Suppose I have a vector

$a = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $b = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$c_1 a + c_2 b$ where c_1, c_2

for $c = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is a linear combination of a and b.

Similarly $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ is also a linear combination of a & b

- Notion can be extended to more than 2 vectors.
- Linear Independence :-

$\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ are linearly independent

However $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ are linearly dependent

Also $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is linearly dependent on $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ & $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$

→ Rank of a Matrix :-

No. of linearly Independent row/column in a matrix is called Rank of matrix

$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ Rank $(A) = 1$

$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ Rank $(A) = 3$

$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ Rank $(A) = 2$

$\boxed{\text{Rank}(A) = m = n}$

→ Basis

Set of LI vectors in $\mathbb{R}^d$ which can represent entire set of linear combination

Eg For $\mathbb{R}^2$ $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ form basis

For $\mathbb{R}^3$ $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ form basis

→ Column Space or Range of A

All vectors that can be represented by linear combination of columns of A

correct thing to matrix vector Multiplication

- Matrix Decomposition

$A = U \Sigma V^T$

$m \times n \quad m \times m \quad m \times n \quad n \times n$

→ Here columns of U & V are orthonormal

Σ has diagonal entries as non negative numbers

Geometric Interpretation
$A = U \Sigma V^T$

$$\boxed{A = \sum_{i=1}^{d} \sigma_i u_i v_i^T}$$

Also $A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$ is best low rank approximation