Linear Algebra

Scalar, Vector, Orthogonal Vector, Matrix Operation

by Dr. Tumpa Banerjee, Dept of MCA

System of linear equation

- Machine learning model is presented as a function and probability distribution of features.
- The objective is learning involves determining the optimal value of the parameter that best fits the model or minimizes the loss function.
- Optimization algorithms deal with maximizing and minimizing an objective defined using linear algebraic equations.

A firm produces three products. These products are processed on three different machines. The time required to manufacture one unit of each of the three products and the daily capacity of the three machines are given in the table below:

Machine	Time per unit (Minutes)			Machine Capacity (minutes/day)
	Product 1	Product 2	Product 3	
M_1	2	3	2	440
M_2	4	-	3	470
M_3	2	5	-	430

It is required to determine the daily number of units to be manufactured for each product. The profit per unit for product 1, 2 and 3 is Rs. 4, Rs.3 and Rs.6 respectively. It is assumed that all the amounts produced are consumed in the market. Formulate the mathematical (L.P.) model that will maximize the daily profit.

Formulation of Linear Programming Model

$$Maximize\ 4x_1 + 3x_2 + 6x_3$$

Subject to

$$2x_1 + 3x_2 + 2x_3 \leq 440$$
$$4x_1 + 0x_2 + 3x_3 \leq 470$$
$$2x_1 + 5x_2 + 0x_3 \leq 430$$
$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$$

Formulation of Linear Programming Model

$$Maximize\ [4,3,6][x_1, x_2, x_3]^T$$

Subject to

$$\begin{bmatrix} 2 & 3 & 2 \\ 4 & 0 & 3 \\ 2 & 5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 440 \\ 470 \\ 430 \end{bmatrix}$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$$

System of linear equation

- Matrices play an important role in machine learning.
- Images and all other inputs, outputs, and parameters of a machine learning model are represented as a matrix.
- Matrix is essential to represent a system of linear equations in compact form.

Scalar and Vector

- A scalar is a numeric value, indicating the magnitude of something. It is a single number having no dimension and is denoted as x, a and a etc.
- A vector is a quantity that represents two things: magnitude and direction.
- It is a one-dimensional array of numbers arranged in order. A vector v can be represented as $(3,4)$

Vector Properties

- The angle between two vectors can be calculated using the dot product formula.
- Let us consider two vectors a and b and the angle between them to be θ. Then, the dot product of two vectors is given by $a.b = |a||b|\cos\theta$.

Vector Properties

Vectors are special objects that can be added together and multiply by scalars and produce another vector.

For any two vectors $v1=\begin{bmatrix} a1 \\ b1 \end{bmatrix}$ and $v2=\begin{bmatrix} a2 \\ b2 \end{bmatrix} \in R^2$, $v1+v2 \in R^2$ and defined as

$$v1+v2 = \begin{bmatrix} a1 \\ b1 \end{bmatrix} + \begin{bmatrix} a2 \\ b2 \end{bmatrix} = \begin{bmatrix} a1+b1 \\ a2+b2 \end{bmatrix}$$

If $v1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $v2 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ then $v1+v2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$

Vector Addition: For any two vector u and $v \in R^n$, $u+v \in R^n$ and

$u+v = v+u$

$(u+v)+w = u+(v+w)$

Additive identity $0 \in R^n$, such that $u+0 = u$

For any $u \in R^n$ there exist an $v \in R^n$ a such that $u+v = 0$

Norms

- A norm on a vector space V is a function $||.||: V \to \mathbb{R}$
- Which assigns each vector x its length $||x|| \in \mathbb{R}$ such that for all $\lambda \in \mathbb{R}$ and $x, y \in V$ the following hold:
- $||\lambda x|| = |\lambda|||x||$
- $||x + y|| \leq ||x|| + ||y||$
- $||x|| \geq 0$ and $||x|| = 0 \Rightarrow x = 0$

Norms

Manhattan Norm :

$$||x||_1 = \sum|x_i|$$

Euclidean Norm:

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

Matrix

- $A_{m\times n} = \begin{pmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}$^$(m \times n)$
- Matrix Addition:
- $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{m\times n}$, then
- $A_{m\times n} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} \cdots & a_{mn} + b_{mn} \end{pmatrix} \in \mathbb{R}^{m\times n}$

- Elementwise matrix multiplication is called Hadamard product

Matrix Multiplication

- For the matrices $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times p}$
- $C_{m\times p} = AB = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}b_{j1} & \sum_{j=1}^{n} a_{1j}b_{j2} \cdots & \sum_{j=1}^{n} a_{1j}b_{jp} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} a_{mj}b_{j1} & \sum_{j=1}^{n} a_{m1j}b_{j2} \cdots & \sum_{j=1}^{n} a_{mj}b_{jp} \end{pmatrix}$

Identity Matrix

- Identity matrix is a square matrix defined as
- $\mathbb{I}_n = a_{ij} = f(x) = \begin{cases} 1, & if\ i = j \\ 0, & if\ i \neq j \end{cases}$

Matrix Operation

- Associative
- $(A+B)+C = A+(B+C)$
- $(AB)C = A(BC)$
- Distributive
- $(A+B)C = AC + BC$
- Multiplication with identity matrix
- $\forall A \in \mathbb{R}^{m\times n}: \mathbb{I}_m A = A\mathbb{I}_n = A$

Matrix Inverse and Transpose

- For a square matrix A. The matrix A is called inverse of A if $AB = I = BA$. The inverse of A is denoted as A^{-1}
- Converting rows to columns and columns to rows, we get transpose of a matrix. Transpose of a matrix A is denoted as A^{-1}.
- A matrix A is called symmetric if $A = A^T$

Matrix Properties

- $AA^{-1} = \mathbb{I} = A^{-1}A$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A+B)^{-1} \neq A^{-1} + B^{-1}$
- $(A^T)^T = A$
- $(AB)^T = B^T A^T$

Inner Product

- Inner products facilitate the introduction of intuitive geometrical concepts, such as the length of a vector and the angle or distance between two vectors.
- A primary function of inner products is to determine whether vectors are orthogonal to each other.

Dot Product and Length of vectors

- A particular type of inner product is a scalar product/dot product and is defined as
- $x^T y = \sum_{i=1}^{n} x_i y_i$
- Length: The length of a vector can be computed using the inner product.
- $||x|| = \sqrt{< x, x >}$

Distance and Metric:

- Consider an inner product space $(V, <.,.>)$. Then
- $d(x, y) = ||x - y|| = \sqrt{< x - y, x - y >}$
- Is called the distance between x and y for $x, y \in V$.
- The mapping $V \times V \to \mathbb{R}$, $(x, y) \to d(x, y)$ is called a metric.

Distance and Metric:

A metric d satisfies the following:

1. d is positive definite i.e. $d(x, y) \geq 0$ for all $x, y \in V$ and $d(x, y) = 0 \Leftrightarrow x = y$

2. d is symmetric i.e. $d(x, y) = d(y, x)$ for all $x, y \in V$

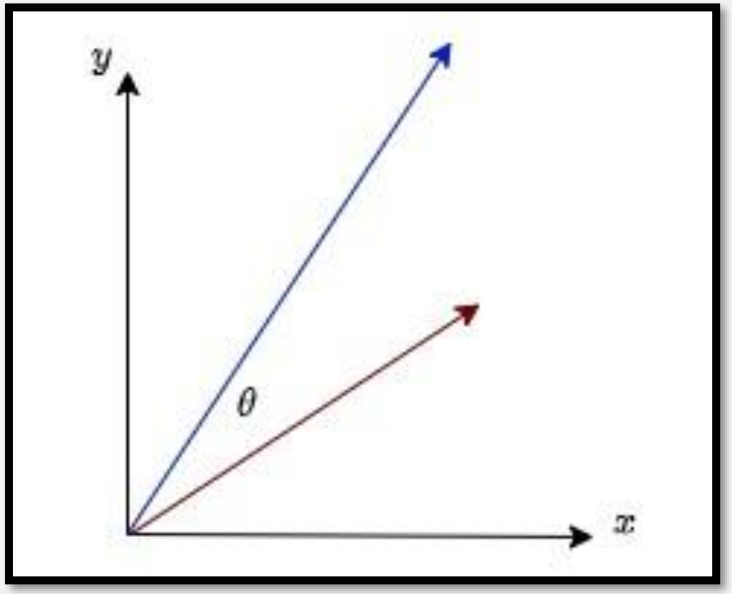
3. Triangular inequality: $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in V$

Cauchy-Schwarz Inequality

- Cauchy-Schwarz Inequality: For an inner product vector space (V,<.,.>) the induced norm ||.|| satisfies the Cauchy-Schwarz inequality
- $|< x, y > | \leq ||x||.||y||$

Angels between two vectors

- The inner product is also used to calculate the angle between two vectors. Using the Cauchy-Schwarz Inequality, we can define the angle θ between two vectors x and y.
- $-1 \leq \frac{|<x,y>|}{||x||||y||} \leq 1$
- There exists an unique $\theta \in [0, \pi]$ with $\cos\theta = \frac{|<x,y>|}{||x||||y||}$



Orthogonal Vector

- Two vectors x and y are orthogonal if and only if $< x, y > = 0$, and we can say $x \perp y$. If additionally $||x|| = 1 = ||y||$, i.e. the vectors are unit vectors, then x and y are orthonormal.