Week-1

Author: Karthik Thiagarajan

1. Data

Data comes in various forms:

- Images
- Videos
- Text
- Audio
- Tabular data

Students	Mathematics	Physics
A	7	6
B	5	3
C	7	9
D	10	8
E	9	7

Each row in this table is called a data-point. Each student is described by a set of features. In this example, the mathematics and physics marks are the features. A feature is something that describes the data-point. The term attribute is an alternative to feature. The number of features in a dataset is denoted by d. The number of data-points is denoted by n.

2. Conventions

The tabular dataset can be described in terms of a data matrix X. The convention we follow in our course is to represent the data-points as the columns. So for the table given above, the data-matrix looks like:

$$X = \begin{bmatrix} 7 & 5 & 7 & 10 & 9 \\ 6 & 3 & 9 & 8 & 7 \end{bmatrix}$$

This is a $d \times n$ matrix. In general, if there are n data-points, where each data-point is in $\mathbb{R}^d$, we have:

$$X = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix}$$

3. Vectors

All vectors in this course are column vectors:

$$x = (1, 2, 3)$$

is represented as:

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

This means that:

$$x^T = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

4. Some Algebra

The following are some important operations on vectors:

4.1. Addition

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ -5 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

4.2. Scalar multiplication

$$3 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

4.3. Dot product

$$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, y = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

$$\begin{aligned} x \cdot y &= x^T y \\ &= \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} \\ &= -1 \end{aligned}$$

4.4. Norm

$$x = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \implies ||x|| = \sqrt{3^2 + 4^2} = 5$$

4.5. Outer product

$$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, y = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

$$\begin{aligned} xy^T &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 3 & -2 \\ 6 & -4 \end{bmatrix} \end{aligned}$$

4.6. Matrix-vector product

$$A = \begin{bmatrix} 1 & 5 \\ 3 & 2 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$$\begin{aligned} Ab &= 2 \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 3 \cdot \begin{bmatrix} 5 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 17 \\ 12 \end{bmatrix} \end{aligned}$$

Important: Matrix-vector multiplication is equivalent to a linear combination of the columns of the matrix, where the multipliers come from the vector.

4.7. Vector-Matrix

$$A = \begin{bmatrix} 1 & 5 \\ 3 & 2 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$$\begin{aligned} b^T A &= 2 \cdot \begin{bmatrix} 1 & 5 \end{bmatrix} + 3 \cdot \begin{bmatrix} 3 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 11 & 16 \end{bmatrix} \end{aligned}$$

4.8. Matrix-Matrix

There are four ways to multiply two matrices:

$$\begin{matrix} A \to m \times n \\ B \to n \times p \end{matrix}, C = AB \to m \times p$$

4.8.1. Method-1 (Matrix-column)

$$B = \begin{bmatrix} | & & | \\ b_1 & \cdots & b_p \\ | & & | \end{bmatrix}$$

$$\begin{aligned} C &= AB \\ &= \begin{bmatrix} | & & | \\ Ab_1 & \cdots & Ab_p \\ | & & | \end{bmatrix} \end{aligned}$$

4.8.2. Method-2 (row-matrix)

$$A = \begin{bmatrix} - & a_1^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}$$

$$\begin{aligned} C &= AB \\ &= \begin{bmatrix} - & a_1^T B & - \\ & \vdots & \\ - & a_m^T B & - \end{bmatrix} \end{aligned}$$

4.8.3. Method-3 (row-column)

$$A = \begin{bmatrix} - & a_1^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}, B = \begin{bmatrix} | & & | \\ b_1 & \cdots & b_p \\ | & & | \end{bmatrix}$$

$$
\begin{aligned}
C &= AB \\
&= \begin{bmatrix} & \vdots & \\ \cdots & a_i^T b_j & \cdots \\ & \vdots & \end{bmatrix}
\end{aligned}
$$

4.8.4. Method-4 (column-row)

$$
A = \begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix}, B = \begin{bmatrix} - & b_1^T & - \\ & \vdots & \\ - & b_n^T & - \end{bmatrix}
$$

$$
\begin{aligned}
C &= AB \\
&= a_1 b_1^T + \cdots + a_n b_n^T
\end{aligned}
$$