Models, Algorithms and Data (MAD): Introduction to Computing

Spring semester 2021

Prof. Dr. Jens Honore Walther
Dr. Georgios Arampatzis
ETH Zentrum, CLT
CH-8092 Zürich

Set 1

Issued: 12.03.2021

Question 1: Sample solution

Here is the suggested solution. We take the year and run time as x and y variables, respectively. To make the calculation simpler, we remove 2000 years so that the table of data is given as below:

x	4	8	12	16
y	10.93	10.78	10.75	10.71

We thus want to predict the time at $x = 20$. Assuming a linear relation between the year and the performance we must find the coefficients w_1 and w_2 in:

$$y = w_1 + w_2 x. \tag{1}$$

We can rewrite the problem in matrix form as:

$$A\mathbf{w} = \mathbf{y}\ , \quad \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \tag{2}$$

The goal is to solve the normal equation:

$$\left(A^T A\right) \mathbf{w} = A^T \mathbf{y} \tag{3}$$

We could simply solve this equation in matrix form. However for this simple linear fit we can rewrite the normal equation as:

$$\begin{bmatrix} N & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i y_i \end{bmatrix} \tag{4}$$

This can be solved as it is a square (2x2) matrix.The solution is given as:

$$
\begin{aligned}
w_1^* &= \frac{\left(\sum_{i=1}^N x_i^2\right)\left(\sum_{i=1}^N y_i\right) - \left(\sum_{i=1}^N x_i\right)\left(\sum_{i=1}^N x_i y_i\right)}{N\left(\sum_{i=1}^N x_i^2\right) - \left(\sum_{i=1}^N x_i\right)^2} \\
w_2^* &= \frac{N\left(\sum_{i=1}^N x_i y_i\right) - \left(\sum_{i=1}^N x_i\right)\left(\sum_{i=1}^N y_i\right)}{N\left(\sum_{i=1}^N x_i^2\right) - \left(\sum_{i=1}^N x_i\right)^2}.
\end{aligned}
\tag{5}
$$

From the lecture, we calculate each component in the table as below.

	x	y	x^2	xy
	4	10.93	16	43.72
	8	10.78	64	86.24
	12	10.75	144	129
	16	10.71	256	171.36
$\sum$	40	43.17	480	430.32

Therefore,

$$
\begin{aligned}
w_2^* &= \frac{N\sum(xy) - \sum x \sum y}{N\sum x^2 - (\sum x)^2} \\
&= \frac{4 \times 430.32 - 40 \times 43.17}{4 \times 480 - 40^2} \\
&= \frac{1721.28 - 1726.8}{1920 - 1600} \\
&= \frac{-5.52}{320} = -0.01725.
\end{aligned}
\tag{6}
$$

Furthermore, we could use Eq. (5) to calculate w_1^*. However, here we directly use the first equation of Eq. (4),

$$
\begin{aligned}
w_1^* &= \frac{\sum y - w_2^* \sum x}{N} \\
&= \frac{43.17 + 0.01725 \times 40}{4} \\
&= \frac{43.86}{4} = 10.965.
\end{aligned}
\tag{7}
$$

The relation is summarized as

$$
y = 10.965 - 0.01725x. \tag{8}
$$

To predict what happens in year $x = 20$, we have

$$
y = 10.965 - 0.01725 \times 20 = 10.62 \tag{9}
$$

Hence the champion in 2020 is expected to run 10.62 seconds for 100 meters.