Linear Model: Least Square Approach

TF4063

Fadjar Fathurrahman

Program Studi Teknik Fisika
Institut Teknologi Bandung

The material in this note is based on Rogers2017.

Simpe Linear Model

Given pair of data (x, t) where x are inputs dan t are targets, a linear model with parameter (w_0, w_1) can be written as:

$$t = f(x; w_0, w_1) = w_0 + w_1 x \tag{1}$$

We want to find model parameters w_0 and w_1 which are best for describing our data.

For n-th data we can write

$$\mathcal{L}_n \equiv (t_n - f(x_n; w_0, w_1))^2 \tag{2}$$

By averaging contributions from all data:

$$\mathcal{L} = \frac{1}{N} \sum_{n=1}^{N} \mathcal{L}_n = \frac{1}{N} \sum_{n=1}^{N} (t_n - f(x_n; w_0, w_1))^2 \tag{3}$$

We will call this quantity as loss function

Simple Linear Model

We can find the parameters (w_0, w_1) by using minimization procedures:

$$\arg\min_{w_0, w_1} \frac{1}{N} \sum_{n=1}^{N} \mathcal{L}_n \tag{4}$$

For our particular case of Eq. (3), we can found this analytically, i.e. calculating the first derivatives of $\mathcal{L}$ with respect to w_0 and w_1, equating them to zero, and solve the resulting equations for w_0 and w_1. For more general cases, we can use various numerical optimization procedures such as gradient descent methods.

We begin by writing our loss function as:

$$\begin{aligned}
\mathcal{L} &= \frac{1}{N} \sum_{n=1}^{N} \left(t_n - (w_0 + w_1 x_n)\right)^2 \\
&= \frac{1}{N} \sum_{n=1}^{N} \left(w_1^2 x_n^2 + 2 w_1 x_n (w_0 - t_n) + w_0^2 - 2 w_0 t_n + t_n^2\right)
\end{aligned}$$

Now we find the first derivatives of $\mathcal{L}$ with respect to w_0, w_1 and equating them to zero.

$$\frac{\partial \mathcal{L}}{\partial w_1} = 2w_1 \frac{1}{N}\left(\sum_{n=1}^{N} x_n^2\right) + \frac{2}{N}\left(\sum_{n=1}^{N} x_n(w_0 - t_n)\right) = 0$$
$$\frac{\partial \mathcal{L}}{\partial w_0} = 2w_0 + 2w_1 \frac{1}{N}\left(\sum_{n=1}^{N} x_n\right) - \frac{2}{N}\left(\sum_{n=1}^{N} t_n\right) = 0$$

We obtain

$$\begin{aligned} w_1 &= \frac{\overline{xt} - \overline{x}\,\overline{t}}{\overline{x^2} - \overline{x}^2} \\ w_0 &= \overline{t} - w_1 \overline{x} \end{aligned} \tag{5}$$

where symbols with overline denotes their average value, for examples

$$\overline{x} = \frac{1}{N}\sum_{n=1}^{N} x_n$$
$$\overline{t} = \frac{1}{N}\sum_{n=1}^{N} t_n$$

$$\mathbf{x}_n = \begin{bmatrix} 1 \\ x_n \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$

$$f(x_n; w_0, w_1) = \mathbf{w}^\mathsf{T}\mathbf{x}_n \tag{6}$$

$$\mathcal{L} = \frac{1}{N}\sum_{n=1}^{N}\left(t_n - \mathbf{w}^\mathsf{T}\mathbf{x}_n\right)^2 \tag{7}$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^\mathsf{T} \\ \mathbf{x}_2^\mathsf{T} \\ \vdots \\ \mathbf{x}_N^\mathsf{T} \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}$$

$$\mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_N \end{bmatrix} \tag{8}$$

Example slide

Column 1

Column 2