01:640:495 - Workshop 15

Pranav Tikkawar

April 14, 2025

1. You have a training dataset $\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$. You believe that given x, y is normally distributed with a mean that is a second-order polynomial in x, and a variance σ^2 that is independent of x. That is,

$$y \sim N(\theta_0 + \theta_1 x + \theta_2 x^2, \sigma^2).$$

In other words,

$$y = \theta_0 + \theta_1 x + \theta_2 x^2 + \text{Noise},$$

where Noise $\sim N(0, \sigma^2)$. Given the training dataset, write the likelihood function, and use it to identify the best parameters using the idea of maximum likelihood estimation. Then, find those parameters for the data set $\{(0, 0), (1, 0), (-1, 1), (1, 2)\}$.

Solution: For this mdoel, we can take our $X = (1, x, x^2)$ and $\Theta = (\theta_0, \theta_1, \theta_2)$. The probability function is given by:

$$P(y|x, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta^T X - y)^2}{2\sigma^2}}$$

And out likelihood function is given by:

$$\Theta_{MLE} = \text{argmax}_\theta \prod_{i=1}^{N} P(y_i | x_i, \theta)$$

$$= \text{argmin}_\theta \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta^T X_i - y_i)^2}{2\sigma^2}}$$

$$= \text{argmin}_\theta \sum_{i=1}^{N} (X\Theta - y_i)^2$$

$$= \text{argmin}_\theta ||X\Theta - y||^2$$

We can see that this is a least squares problem. So this with the data of $y = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}$

and $X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. We can solve for Θ by solving

$$\Theta = (X^T X)^{-1} X^T y$$

$$\Theta = \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \right)^{-1} \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right)$$

$$= \left(\begin{bmatrix} 4 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{bmatrix} \right)^{-1} \left(\begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix} \right)$$

$$= \begin{bmatrix} 1 & 0 & -1 \\ 0 & \frac{3}{8} & -\frac{1}{8} \\ -1 & -\frac{1}{8} & \frac{11}{8} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Calculating this gives us $\Theta = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.