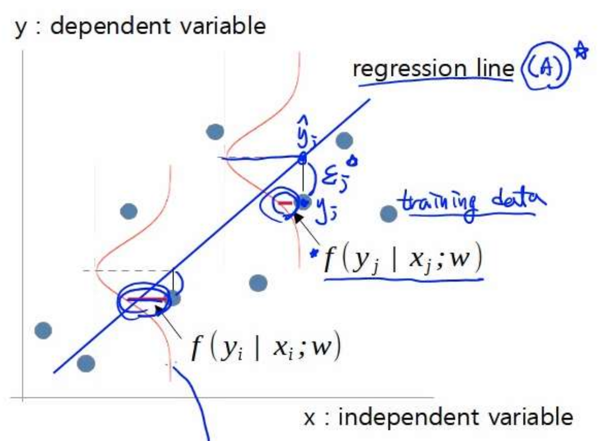
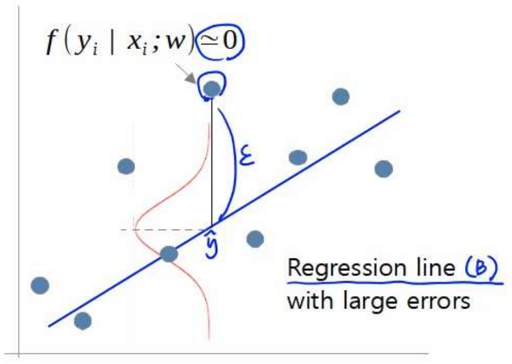
MLE

Sunday 19 May 2024 2:12 AM

Estimating a regression line using the Maximum Likelihood Estimates (MLE)

- Suppose we know the data consists of values drawn from $(y=wx+b+\varepsilon)$ ε is independent, identically distributed, and normally distributed.
- What are the parameter values w and b for which the observed data have the greatest probability?
- This can be solved with MLE, and the result is the same as the Least Squares Method.





$$\hat{y}_i = w_1 x_{1,i} + w_2 x_{2,i} + \ldots + w_k x_{k,i} + b$$

$\varepsilon \sim N(0, \sigma^2)$

$$f(y_i \mid x_i; w, b) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp \frac{-(\overbrace{y_i - \hat{y}_i}^{\varepsilon})^2}{2\sigma^2}$$

$$\max_{w,b} L(w,b) = \prod_{i=1}^{N} f(y_i \mid x_i; w, b) \quad \leftarrow \text{likelihood function}$$

log likelihood function

$$\log(L(w,b)) = \sum_{i=1}^{N} \log(f(y_i \mid x_i; w, b))$$

$$\max_{w,b} \log(L(w,b)) = N \log \cancel{\frac{1}{\sqrt{2\pi}\,\sigma}} - \frac{1}{2\sigma^2} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

- Maximize log likelihood

$$\max_{w,b} \log(L(w,b)) = \max_{w,b} \left(-\sum_i (y_i - \hat{y}_i)^2\right)$$

- Minimize squared error
  = least squares method

$$\max_{w,b} \left(-\sum_i (y_i - \hat{y}_i)^2\right) \rightarrow \min_{w,b} \sum_i (y_i - \hat{y}_i)^2$$

- Maximizing the log-likelihood is equivalent to minimizing the squared error.