Maximum Likelihood Estimate (Linear Regression)

(Consider the 1D case)

$$C(\theta) = P(y|X;\theta)$$

$$= \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y^{(i)} - \theta^T x^{(i)})^2}{2\sigma^2}\right)$$

Likelihood and probability are similar. However when we say likelihood we say that y and X in $P(y|X,\theta)$ are fixed.

Likelihood of parameters vs. Probability of data.

Maximum Likelihood:

$$\theta_{ML} = \underset{\theta}{\arg\max} \prod_{i=1}^{M} P(y^{(i)}|X;\theta,\sigma)$$

We do not treat θ as a random variable.
This is frequentist approach.
We are trying to estimate some true value of θ.

Choose θ to maximize $\prod_{i=1}^{N} P(y^{(i)}|X;\theta,\sigma)$

Instead lets maximize $\log \prod_{i=1}^{M} P(y^{(i)}|X;\theta,\sigma)$

$$= \sum_{i=1}^{M} \log P(y^{(i)}|X;\theta,\sigma)$$

$\log ab$
$= \log a + \log b$

$$= \sum_{i=1}^{m} \log \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y^{(i)} - \theta^T x^{(i)})^2}{2\sigma^2}\right)$$

$$= \sum_{i=1}^{m} \log \frac{1}{\sqrt{2\pi}\sigma} + \sum_{i=1}^{m} \log \exp\left(-\frac{(y^{(i)} - \theta^T x^{(i)})^2}{2\sigma^2}\right)$$

$$= m \log \frac{1}{\sqrt{2\pi}\sigma} + \sum_{i=1}^{m} \left\{ -\left(\frac{(y^{(i)} - \theta^T x^{(i)})^2}{2\sigma^2}\right) \right\}$$

$\therefore$ To maximize $\ell(\theta)$, we need to minimize $\sum_{i=1}^{m} \frac{(y^{(i)} - \theta^T x^{(i)})^2}{2\sigma^2}$ which is the same as the least-square error.

θ, it seems, doesn't depend upon σ^2.