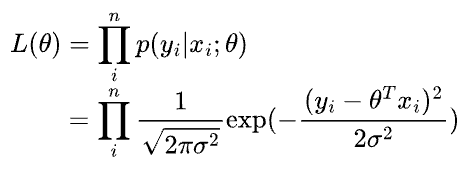
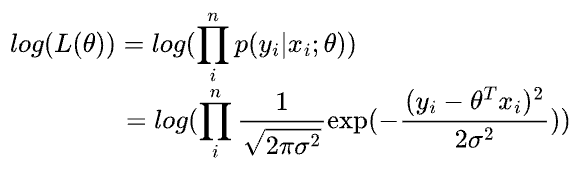
**maximum likelihood estimation (MLE**

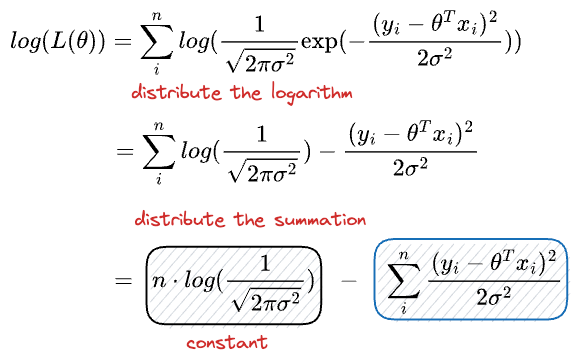
https://www.dailydoseofds.com/content/images/2023/08/image-27.pngLikelihood function

Likelihood function

Since the log transformation is monotonic, we use the log-likelihood below to optimize MLE.

Log-likelihood function

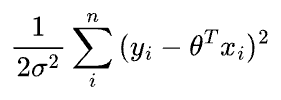
Simplifying, we get:

Simplifying log-likelihood function

To find the best Gaussian that describes the true underlying model which generates our data, in other words, the best θ, we need to find the parameters that maximizes the log-likelihood.

**Maximizing the above expression is equivalent to minimizing the second term.**

**And if you notice closely, it's precisely the least squares.**

Squared Error