Maximum Likelihood Estimation

Mustaf Ahmed

April 9, 2025

1 Maximizing log-likelihood

In statistical inference, Maximum Likelihood Estimation (MLE) is a method of producing a model that best accounts for the observed data by maximizing the likelihood of the observed data; In this way, we select a model that best predicts that data we observed.

Optimizing the likelihood function becomes easier when we instead maximize the log-likelihood. We'll justify this alternative optimization in what follows.

Theorem 1 (Equivalence of Likelihood and Log-Likelihood Maximizers)**.** *Let $L(\theta|\mathbf{x})$ be the likelihood function for parameters θ given data $\mathbf{x}$, assumed to be positive. Let $\log L(\theta|\mathbf{x})$ be the corresponding log-likelihood function. A parameter value $\hat{\theta}$ maximizes the likelihood function if and only if it also maximizes the log-likelihood function. Formally:*

$$\hat{\theta} = \arg\max_{\theta} L(\theta|\mathbf{x}) \iff \hat{\theta} = \arg\max_{\theta} \log L(\theta|\mathbf{x})$$

Proof. For the forward direction, assume that $\hat{\theta} = \arg\max_{\theta} L(\theta|\mathbf{x})$. By the definition of argument maximization, we have $L(\theta^{'}|\mathbf{x}) \leq L(\hat{\theta}|\mathbf{x})$. Since $\log(x)$ is strictly increasing, $\log(L(\theta^{'}|\mathbf{x})) \leq \log(L(\hat{\theta}|\mathbf{x}))$, so $\hat{\theta} = \arg\max_{\theta} \log L(\theta|\mathbf{x})$.

For the backwards direction, assume that $\hat{\theta} = \arg\max_{\theta} \log L(\theta|\mathbf{x})$. By the definition of argument maximization, we have $\log(L(\theta^{'}|\mathbf{x})) \leq \log(L(\hat{\theta}|\mathbf{x}))$. Since e^x is strictly increasing, we have

$$e^{\log(L(\theta^{'}|\mathbf{x}))} \leq e^{\log(L(\hat{\theta}|\mathbf{x}))}$$

, and by simplifying we get

$$L(\theta^{'}|\mathbf{x}) \leq L(\hat{\theta}|\mathbf{x})$$

$\square$