GMM Appendix A: MLE for Gaussian Parameters: Complete Derivation

In this appendix, we will derive the maximum likelihood estimates for the parameters of a Gaussian distribution.[1]

1. Courtesy of Claude.ai

Setup

We have n i.i.d. observations $x_1, x_2, \ldots, x_n$ from $\mathcal{N}(\mu, \sigma^2)$.

The log-likelihood is:

$$\ell(\mu, \sigma^2) = \sum_{i=1}^{n} \log p(x_i \mid \mu, \sigma^2)$$

$$= \sum_{i=1}^{n} \left[-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) - \frac{(x_i - \mu)^2}{2\sigma^2} \right]$$

$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

Derivative with Respect to μ

Taking the Partial Derivative

$$\frac{\partial \ell}{\partial \mu} = \frac{\partial}{\partial \mu}\left[-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right]$$

The first two terms don't depend on μ, so their derivatives are zero:

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{2\sigma^2}\frac{\partial}{\partial \mu}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$= -\frac{1}{2\sigma^2}\sum_{i=1}^{n}\frac{\partial}{\partial \mu}(x_i - \mu)^2$$

Using the chain rule:

Derivative with Respect to σ^2

Taking the Partial Derivative

Let's work with $v = \sigma^2$ for clarity:

$$\ell(\mu, v) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(v) - \frac{1}{2v}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{\partial \ell}{\partial v} = \frac{\partial}{\partial v}\left[-\frac{n}{2}\log(v)\right] + \frac{\partial}{\partial v}\left[-\frac{1}{2v}\sum_{i=1}^{n}(x_i - \mu)^2\right]$$

First Term

$$\frac{\partial}{\partial v}\left[-\frac{n}{2}\log(v)\right] = -\frac{n}{2} \cdot \frac{1}{v} = -\frac{n}{2v}$$

Summary of Results

Log-Likelihood

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

Partial Derivatives

$$\boxed{\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu)}$$

$$\boxed{\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2}$$

Maximum Likelihood Estimates