Statistical signal processing
5CTA0

Estimation theory - Maximum likelihood estimator

Maximum likelihood estimator

- Most popular approach to find an estimator
- If efficient estimator exists, it is the maximum MLE
- Asymptotically unbiased and efficient
- Maximum likelihood estimate $\hat{\theta}_{\mathsf{ML}}$:

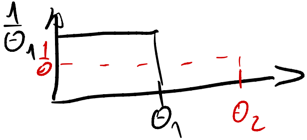
$$\hat{\theta}_{\mathsf{ML}} = \arg\max_{\theta} p(\mathbf{x}; \theta)$$

  for fixed $\mathbf{x}$
- Remark: Maximum might not exist or might not be unique

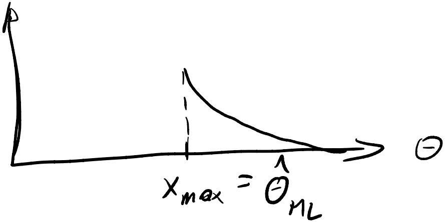
Example I

Let $\{x_0, x_1, \ldots, x_{N-1}\}$ be IID uniformly distributed random variables with PDF

$$p(x_i;\theta) = \begin{cases} \theta^{-1} & 0 \le x_i \le \theta, \\ 0 & \text{else.} \end{cases}$$



$$p(\mathbf{x};\theta) = \begin{cases} \theta^{-N}, & \text{if } 0 < x_i < \theta \text{ for } 0 \le i \le N-1 \\ 0, & \text{else.} \end{cases}$$



Finding the MLE

- If $p(\mathbf{x};\theta)$ is continuously differentiable and if the maximum is interior to the range of θ, then the MLE can be found by equating the derivative of $p(\mathbf{x};\theta)$ with respect to θ to zero and solving for θ:

$$\frac{\partial}{\partial\theta}p(\mathbf{x};\theta) = 0$$

- Attention: Necessary but not sufficient
- Logarithm is strictly monotonic:

$$\arg\max_{\theta}\ p(\mathbf{x};\theta) = \arg\max_{\theta}\ \ln p(\mathbf{x};\theta)$$

- Likelihood equation:

$$\frac{\partial}{\partial\theta}\ln p(\mathbf{x};\theta) = 0$$

Example II: Estimation of a DC voltage

$$x_n = A + w_n, \qquad w_n \sim \mathcal{N}(0, \sigma^2)$$

$$p(\mathbf{x};\theta) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x_n - A)^2\right)$$

$$\frac{\partial}{\partial A}\ln p(\mathbf{x};A) = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}(x_n - A) = \frac{1}{\sigma^2}\left(\sum_{n=0}^{N-1} x_n - NA\right) = 0 \quad \Big| \cdot \sigma^2$$

$$\hat{A}_{ML} = \frac{1}{N}\sum_{n=0}^{N-1} x_n$$

MLE - Efficiency

$$\frac{\partial}{\partial \theta} \ln p(\underline{x}, \theta) = \underbrace{I(\theta)}_{>0} \underbrace{(g(\underline{x}) - \theta)}_{0} = 0$$

$$g(x) = \hat{\theta}_{ML}$$

MLE - Properties

- Consistent:

$$\lim_{N\to\infty} \Pr\left[|g(\mathbf{x}) - \theta| > \varepsilon\right] = 0$$

$\mathcal{E} \in \mathbb{R}^{+}$

- Asymptotically unbiased:

$$\mathrm{E}\left[\hat{\theta}_{\mathsf{ML}}\right] \to \theta$$

- Asymptotically efficient:

$$\mathrm{Var}\left[\hat{\theta}_{\mathsf{ML}}\right] \to \mathcal{I}^{-1}(\theta),$$

- Asymptotically normal distributed:

$$\hat{\theta}_{\mathsf{ML}} \sim \mathcal{N}\left(\theta, \mathcal{I}^{-1}(\theta)\right).$$