Generalized Likelihood Ratio Test

Jessica Fridrich

EECE 566

Introduction

$$\begin{aligned} H_0 : &\quad \mathbf{x} \sim p(\mathbf{x}; \boldsymbol{\theta}_0) \\ H_1 : &\quad \mathbf{x} \sim p(\mathbf{x}; \boldsymbol{\theta}_1) \end{aligned}$$

- $\boldsymbol{\theta}_0$ (resp. $\boldsymbol{\theta}_1$) is the parameter vector under H_0 (resp. under H_1)
- $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$ may be different parameters
- Previous lecture $=$ Bayesian approach (assign prior pdf's to the parameters under both H_0 and H_1, $p(\boldsymbol{\theta}_0)$ and $p(\boldsymbol{\theta}_1)$, and integrate)

Generalized likelihood ratio test (GLRT)

- Convert composite test to simple by plugging into the pdf's MLE estimates of the parameters
- Test statistic

$$\text{(GLRT): } L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_1, H_1)}{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_0, H_0)} \underset{1}{\overset{0}{\lessgtr}} \gamma$$

- $\hat{\boldsymbol{\theta}}_i$ is the maximum likelihood estimate (MLE) of the parameters under H_i
- MLE: $\hat{\boldsymbol{\theta}}_i = \text{argmax}_{\boldsymbol{\theta}} p(\mathbf{x}; \boldsymbol{\theta}, H_i)$, maximizes the likelihood under H_i

Generalized likelihood ratio test (GLRT)

- GLRT can also be written as

$$\text{(GLRT): } L_G(\mathbf{x}) = \frac{\max_{\boldsymbol{\theta}_1} p(\mathbf{x}; \boldsymbol{\theta}_1, H_1)}{\max_{\boldsymbol{\theta}_0} p(\mathbf{x}; \boldsymbol{\theta}_0, H_0)} \underset{1}{\overset{0}{\lessgtr}} \gamma$$

- Suppose pdf under H_0 is completely known

$$\text{(GLRT): } L_G(\mathbf{x}) = \frac{\max_{\boldsymbol{\theta}_1} p(\mathbf{x}; \boldsymbol{\theta}_1, H_1)}{p(\mathbf{x}; H_0)} \underset{1}{\overset{0}{\lessgtr}} \gamma$$

$$L_G(\mathbf{x}) = \max_{\boldsymbol{\theta}_1} \frac{p(\mathbf{x}; \boldsymbol{\theta}_1, H_1)}{p(\mathbf{x}; H_0)} \underset{1}{\overset{0}{\lessgtr}} \gamma$$

$$L_G(\mathbf{x}) = \max_{\boldsymbol{\theta}_1} L(\mathbf{x}; \boldsymbol{\theta}_1) \underset{1}{\overset{0}{\lessgtr}} \gamma$$

- Over all possible parameters $\boldsymbol{\theta}_1$ use the largest likelihood ratio test as test statistic.

Example: Detection of unknown DC level in WGN (GLRT)

$$H_0 : \quad x_i = \xi_i$$
$$H_1 : \quad x_i = A + \xi_i$$

$\xi_i \sim \mathcal{N}(0, \sigma^2)$, iid with **known** σ^2

A is **unknown**, two-sided test

- MLE estimator of A:

  under $H_1 : x_i \quad \sim \quad \mathcal{N}(A, \sigma^2)$

- From lecture on MLE: $\hat{A} = \overline{\mathbf{x}}$, the sample mean

Example: Detection of unknown DC level in WGN (GLRT)

$$\text{(GLRT)}: L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{A}, H_1)}{p(\mathbf{x}; H_0)} \underset{1}{\overset{0}{\lessgtr}} \gamma$$

$$\frac{\cancel{(2\pi\sigma^2)^{-n/2}} \exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^n (x_i - \overline{\mathbf{x}})^2\right)}{\cancel{(2\pi\sigma^2)^{-n/2}} \exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^n x_i^2\right)} \underset{1}{\overset{0}{\lessgtr}} \gamma$$

$$\ln L_G(\mathbf{x}) = \frac{-1}{2\sigma^2}\left(\cancel{\sum_{i=1}^n x_i^2} - 2\overline{\mathbf{x}}\sum_{i=1}^n x_i + n\overline{\mathbf{x}}^2 - \cancel{\sum_{i=1}^n x_i^2}\right) \underset{1}{\overset{0}{\lessgtr}} \ln\gamma$$

$$\frac{-1}{2\sigma^2}\left(-2n\overline{\mathbf{x}}^2 + n\overline{\mathbf{x}}^2\right) = \frac{n\overline{\mathbf{x}}^2}{2\sigma^2} \underset{1}{\overset{0}{\lessgtr}} \ln\gamma$$

Example: Detection of unknown DC level in WGN (GLRT)

$$2 \ln L_G(\mathbf{x}) = \frac{n\overline{\mathbf{x}}^2}{\sigma^2} \underset{1}{\overset{0}{\lessgtr}} \ln \gamma$$

$$|\overline{\mathbf{x}}| \underset{1}{\overset{0}{\lessgtr}} \sqrt{\frac{2\sigma^2}{n} \ln \gamma} = \gamma'$$

- This is the ad-hoc detector studied before
- We know its performance and compared it to clairvoyant detector

Example: Detection of **unknown** DC level in WGN of unknown variance (GLRT)

$$
\begin{aligned}
H_0 : &\quad x_i = \xi_i \\
H_1 : &\quad x_i = A + \xi_i
\end{aligned}
$$

$\xi_i \sim \mathcal{N}(0, \sigma^2)$, iid, σ^2 is **unknown**

A is **unknown**, two-sided test

Example: Detection of unknown DC level in WGN of unknown variance (GLRT)

$$\text{(GLRT): } L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{A}, \hat{\sigma}_1, H_1)}{p(\mathbf{x}; \hat{\sigma}_0, H_0)} \underset{1}{\overset{0}{\lessgtr}} \gamma$$

- $\hat{A}, \hat{\sigma}_1$ MLE estimates under H_1
- $\hat{\sigma}_0$ MLE estimate under H_0

Example: Detection of unknown DC level in WGN of unknown variance (GLRT)

- Under H_0

$$x_i \sim \mathcal{N}(0, \sigma^2)$$

- MLE estimator of σ^2 is:

$$\hat{\sigma}_0^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2$$

Example: Detection of unknown DC level in WGN of unknown variance (GLRT)

$$
\begin{aligned}
p(\mathbf{x}; \hat{\sigma}_0^2, H_0) &= (2\pi\hat{\sigma}_0^2)^{-n/2} \exp\left(\frac{-1 \sum_{i=1}^{n} x_i^2}{2\hat{\sigma}_0^2} \right) \\
&= (2\pi\hat{\sigma}_0^2)^{-n/2} \exp\left(\frac{-1 \cancel{\sum_{i=1}^{n} x_i^2}}{2\frac{1}{n}\cancel{\sum_{i=1}^{n} x_i^2}} \right) \\
p(\mathbf{x}; \hat{\sigma}_0^2, H_0) &= (2\pi\hat{\sigma}_0^2)^{-n/2} \exp\left(-n/2\right)
\end{aligned}
$$

Example: Detection of unknown DC level in WGN of unknown variance (GLRT)

- Under H_1

$$x_i \quad \sim \quad \mathcal{N}(A, \sigma^2)$$

- MLE estimator of A is:

$$\hat{A} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{\mathbf{x}}$$

- MLE estimator of σ^2 is:

$$\hat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{\mathbf{x}})^2$$

Example: Detection of unknown DC level in WGN of unknown variance (GLRT)

$$
\begin{aligned}
p(\mathbf{x}; \hat{A}, \hat{\sigma}_1^2, H_1) &= (2\pi\hat{\sigma}_1^2)^{-n/2} \exp\left(\frac{-1 \sum_{i=1}^{n} (x_i - \hat{A})^2}{2\hat{\sigma}_1^2} \right) \\
&= (2\pi\hat{\sigma}_1^2)^{-n/2} \exp\left(\frac{-1 \cancel{\sum_{i=1}^{n} (x_i - \bar{\mathbf{x}})^2}}{2 \frac{1}{n} \cancel{\sum_{i=1}^{n} (x_i - \bar{\mathbf{x}})^2}} \right) \\
p(\mathbf{x}; \hat{A}, \hat{\sigma}_1^2, H_1) &= (2\pi\hat{\sigma}_1^2)^{-n/2} \exp\left(-n/2\right)
\end{aligned}
$$

Example: Detection of unknown DC level in WGN of unknown variance (GLRT)

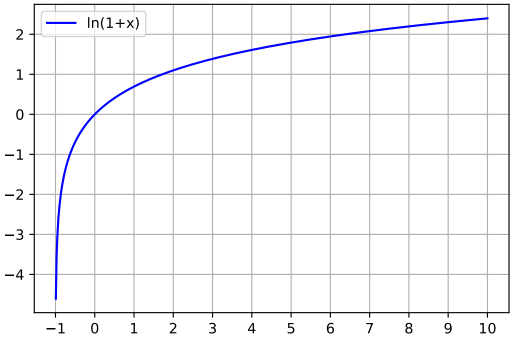
Note that:

$$\begin{aligned}
\hat{\sigma}_0^2 &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{\mathbf{x}} + \overline{\mathbf{x}})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[(x_i - \overline{\mathbf{x}})^2 + 2\overline{\mathbf{x}}(x_i - \overline{\mathbf{x}}) + \overline{\mathbf{x}}^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{\mathbf{x}})^2 + 2\overline{\mathbf{x}}\frac{1}{n}\left[\sum_{i=1}^{n} x_i - n\overline{\mathbf{x}}\right] + \frac{n}{n}\overline{\mathbf{x}}^2 \\
\hat{\sigma}_0^2 &= \hat{\sigma}_1^2 + \overline{\mathbf{x}}^2
\end{aligned}$$

Example: Detection of unknown DC level in WGN of unknown variance (GLRT)

$$
\begin{aligned}
L_G(\mathbf{x}) &= \frac{p(\mathbf{x}; \hat{A}, \hat{\sigma}_1, H_1)}{p(\mathbf{x}|\hat{\sigma}_0, H_0)} \\
&= \frac{(2\pi\hat{\sigma}_1^2)^{-n/2}}{(2\pi\hat{\sigma}_0^2)^{-n/2}} \\
2\ln L_G(\mathbf{x}) &= n\ln\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2} \\
&= n\ln\frac{\hat{\sigma}_1^2 + \overline{\mathbf{x}}^2}{\hat{\sigma}_1^2} \\
2\ln L_G(\mathbf{x}) &= n\ln\left(1 + \frac{\overline{\mathbf{x}}^2}{\hat{\sigma}_1^2}\right)
\end{aligned}
$$

Example: Detection of **unknown** DC level in WGN of unknown variance (GLRT)



Example: Detection of unknown DC level in WGN of unknown variance (GLRT)

- $\ln(1+x)$ is monotone in $x \Rightarrow$ an equivalent test statistic is:

$$T(\mathbf{x}) = \frac{\overline{\mathbf{x}}^2}{\hat{\sigma}_1^2}$$

- We say that the GLRT has normalized the test statistic by $\hat{\sigma}_1^2$ (compared to the known σ^2 case) to allow the threshold to be determined

Example: Detection of unknown DC level in WGN of unknown variance (GLRT)

- Note that we can write: $\xi_i = \sigma u_i$ where $u_i \sim \mathcal{N}(0,1)$

$$
\begin{aligned}
\text{under } H_0 : T(\mathbf{x}) &= \frac{\overline{\mathbf{x}}^2}{\hat{\sigma}_1^2} \\
&= \frac{\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)^2}{\frac{1}{n}\sum_{i=1}^{n}(\xi_i - \overline{\boldsymbol{\xi}})^2} \\
&= \frac{\left(\sigma\sum_{i=1}^{n}u_i\right)^2}{\sum_{i=1}^{n}(\sigma u_i - \sigma/n\sum_{i=1}^{n}u_i)^2} \\
T(\mathbf{x}) &= \frac{\overline{\mathbf{u}}^2}{\sum_{i=1}^{n}(u_i - \overline{\mathbf{u}})^2}
\end{aligned}
$$

- Under H_0: $T(\mathbf{x})$ has a pdf that does not depend on σ^2 $\implies$ we can set the threshold for a given P_{FA}, albeit only numerically
- For large n: Large Data Records (LDR), asymptotically $T(\mathbf{x})$ has a closed-form pdf