Detection Theory

Chapter 9 Unknown Noise Parameters

Marco Munguía M.

December 7th 2010

Outline

- Motivation
- General Considerations
- Hierarchy of detections problems in presence of unknown parameters
  - White Gaussian Noise
    - Known Deterministic Signal
    - Random Signal with Known PDF
    - Deterministic Signal with Unknown Parameters
    - Example 9.3: Signal with Unknown Linear Parameters in WGN with Unknown Variance
  - Colored WSS Gaussian Noise
    - Known Deterministic Signal
    - Rao Test for the General Linear Model with Unknown Noise Parameters

Motivation

- The detection of signals that are not completely known were presented in Ch. 7 and Ch. 8 (deterministic and random signal respectively). However, the PDF of the noise was known.
- The problem of detecting a signal in Gaussian noise when the PDF of the noise is not completely known is of considerable practical interest.
- In this chapter, we explore a number of detectors that are applicable to this problem.

General Considerations

- NP detector:

$$P_{FA} = \Pr\left\{T(x) > \gamma'; \mathrm{H}_0\right\}$$
$$= \int_{\gamma'}^{\infty} p\left(T; \mathrm{H}_0\right) dT$$

- However, when the T(x) is not completely known, then the threshold cannot be determine. One possible approach is to estimate σ^2, assuming that H_0 is true. This approach is refferred to as an *estimate and plug* detector.

- Detection of known DC level (with A > 0) in WGN

$$\gamma' = \sqrt{\hat{\sigma}^2 / N}\, Q^{-1}\left(P_{FA}\right) \quad where \quad \hat{\sigma}^2 = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n]$$

General Considerations cont.

- This estimator is biased when the signal is present: increase the threshold and thereby reduce P_D.

- To avoid the effect of the signal induced bias, one can use additional or reference data samples, known to consist of noise only, and independent of the signal and noise data to form

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{n=0}^{N-1} w_R^2[n] \quad and \quad T\left(x, w_R^2\right) = \frac{\frac{1}{N}\sum_{n=0}^{N-1} x[n]}{\sqrt{\hat{\sigma}_2 / N}}$$

Example 9.1: GLRT for DC Level

$$\frac{p(\mathbf{x}; \hat{\sigma_1^2}, \mathcal{H}_1)}{p(\mathbf{x}; \hat{\sigma_0^2}, \mathcal{H}_0)} > \gamma$$

and

$$\hat{\sigma_0^2} = \frac{1}{N}\sum_{n=0}^{N-1} x^2[n]$$

$$\hat{\sigma_1^2} = \frac{1}{N}\sum_{n=0}^{N-1} (x[n]-A)^2$$

$$L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{\sigma_1^2}, \mathcal{H}_1)}{p(\mathbf{x}; \hat{\sigma_0^2}, \mathcal{H}_0)}$$

$$= \frac{\frac{1}{(2\pi\hat{\sigma_1^2})^{\frac{N}{2}}}\exp\left[-\frac{1}{2\hat{\sigma_1^2}}\sum_{n=0}^{N-1}(x[n]-A)^2\right]}{\frac{1}{(2\pi\hat{\sigma_0^2})^{\frac{N}{2}}}\exp\left[-\frac{1}{2\hat{\sigma_0^2}}\sum_{n=0}^{N-1}x^2[n]\right]}$$

Decide H_1 if:

$$\frac{\hat{\sigma_0^2}}{\hat{\sigma_1^2}} > \gamma^{\frac{2}{N}}.$$

Example 9.1: GLRT for DC Level (cont.)

Equivalent T(x),that is slightly more intuitive is

$$T(\mathbf{x}) = \frac{1}{2A}\left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2} - 1\right) = \frac{\bar{x} - A/2}{\frac{1}{N}\sum_{n=0}^{N-1}(x[n]-A)^2}$$

For weak signals and N → ∞

$$T(\mathbf{x}) \overset{a}{\sim} \begin{cases} \mathcal{N}\left(-\frac{A}{2\sigma^2}, \frac{1}{N\sigma^2}\right) & \text{under } \mathcal{H}_0 \\ \mathcal{N}\left(\frac{A}{2\sigma^2}, \frac{1}{N\sigma^2}\right) & \text{under } \mathcal{H}_1 \end{cases}$$

- No CFAR detector (even asymptotically)
- The GLRT is not found by maximazing the LRT.

Hierarchy of detections problems

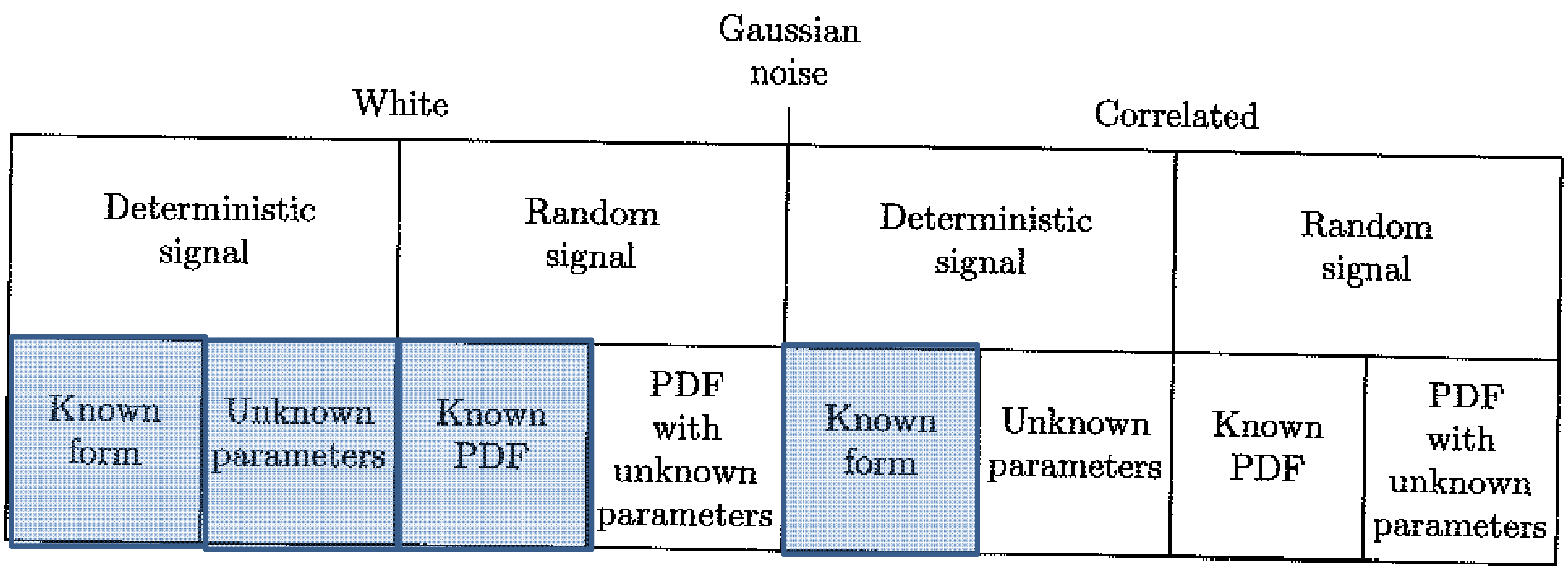


Figure 9.1. Hierarchy of detection problems in presence of unknown parameters.

White Gaussian Noise

- The only possible unknown parameter is the noise variance σ^2.
- The σ^2 needs to be estimated under both hypothesis.
- Under H_0, the MLE of σ^2 is

$$\hat{\sigma_0^2} = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n]$$

and thus

$$p(\mathbf{x}; \hat{\sigma_0^2}, \mathcal{H}_0) = \frac{1}{(2\pi \hat{\sigma_0^2})^{\frac{N}{2}}} \exp\left(-\frac{N}{2}\right)$$

- Under H1, the MLE of σ^2 will depend on the signal assumptions.

Known Deterministic Signal

- The detection problem is

$$\begin{array}{ll} \mathcal{H}_0 : x[n] = w[n] & n = 0, 1, \ldots, N-1 \\ \mathcal{H}_1 : x[n] = s[n] + w[n] & n = 0, 1, \ldots, N-1 \end{array}$$

- The GLRT decides H_1 if

$$L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{\sigma_1^2}, \mathcal{H}_1)}{p(\mathbf{x}; \hat{\sigma_0^2}, \mathcal{H}_0)} > \gamma$$

- The MLE of σ^2 under H_1 is (see page 183, Kay-Vol I)

$$\hat{\sigma_1^2} = \frac{1}{N} \sum_{n=0}^{N-1} (x[n] - s[n])^2$$

and

$$L_G(\mathbf{x}) = \left(\frac{\hat{\sigma_0^2}}{\hat{\sigma_1^2}} \right)^{\frac{N}{2}}.$$

Known Deterministic Signal (cont.)

- If we consider the equivalent test statistic

$$\begin{aligned} T(\mathbf{x}) &= \frac{N}{2}\left(L_G(\mathbf{x})^{\frac{2}{N}} - 1\right) \\ &= \frac{N}{2}\frac{\hat{\sigma_0^2} - \hat{\sigma_1^2}}{\hat{\sigma_1^2}} \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} T(\mathbf{x}) &= \frac{\sum_{n=0}^{N-1} x^2[n] - \sum_{n=0}^{N-1}(x[n]-s[n])^2}{2\hat{\sigma_1^2}} \\ &= \frac{\sum_{n=0}^{N-1} x[n]s[n] - \frac{1}{2}\sum_{n=0}^{N-1} s^2[n]}{\hat{\sigma_1^2}}. \end{aligned}$$

- This dectector is not CFAR

$$T(\mathbf{x}) \stackrel{a}{\sim} \begin{cases} \mathcal{N}\left(-\frac{\mathcal{E}}{2\sigma^2}, \frac{\mathcal{E}}{\sigma^2}\right) & \text{under } \mathcal{H}_0 \\ \mathcal{N}\left(\frac{\mathcal{E}}{2\sigma^2}, \frac{\mathcal{E}}{\sigma^2}\right) & \text{under } \mathcal{H}_1 \end{cases}$$

Random Signal with Known PDF

- Now, it is assumed that s[n] is a Gaussian random process with known PDF: zero mean and covariance matrix $\mathbf{C}_s$
- The PDF under H_1 is

$$p(\mathbf{x};\sigma^2,\mathcal{H}_1)=\frac{1}{(2\pi)^{\frac{N}{2}}\det^{\frac{1}{2}}(\mathbf{C}_s+\sigma^2\mathbf{I})}\exp\left[-\frac{1}{2}\mathbf{x}^T(\mathbf{C}_s+\sigma^2\mathbf{I})^{-1}\mathbf{x}\right]$$

- To find the MLE of σ^2 under H_1, an equivalent expression of the PDF based on eigendecomposition of $\mathbf{C}_s$ is used

$$\ln p(\mathbf{x};\sigma^2,\mathcal{H}_1)=-\frac{N}{2}\ln 2\pi-\frac{1}{2}\sum_{i=1}^{N}\left[\ln(\lambda_{s_i}+\sigma^2)+\frac{(\mathbf{v}_i^T\mathbf{x})^2}{\lambda_{s_i}+\sigma^2}\right]$$

Random Signal with Known PDF (cont.)

- The MLE of σ^2 is found by minimazing

$$J(\sigma^2) = \sum_{i=1}^{N} \left[\ln(\lambda_{s_i} + \sigma^2) + \frac{(\mathbf{v}_i^T \mathbf{x})^2}{\lambda_{s_i} + \sigma^2} \right]$$

- This cannot be solved analytically. Let's consider the case when the signal is very weak or $\lambda_{s_i} << \sigma^2$

$$\hat{\sigma}_1^2 = \max\left(0, \frac{1}{N} \sum_{n=0}^{N-1} x^2[n] - \frac{1}{N} \mathrm{tr}(\mathbf{C}_s)\right)$$

Random Signal with Known PDF (cont.)

- Recall that we are maximizing the PDF over $0 < \sigma^2 < \infty$

$$2 \ln L_G(\mathbf{x}) = \sum_{i=1}^{N} \left[\ln \frac{\hat{\sigma_0^2}}{\lambda_{s_i} + \hat{\sigma_1^2}} - \frac{(\mathbf{v}_i^T \mathbf{x})^2}{\lambda_{s_i} + \hat{\sigma_1^2}} + 1 \right]$$

- Finally, we decide H_1 if

$$T(\mathbf{x}) = \sum_{i=1}^{N} \left[\ln \frac{\hat{\sigma_0^2}}{\lambda_{s_i} + \hat{\sigma_1^2}} - \frac{(\mathbf{v}_i^T \mathbf{x})^2}{\lambda_{s_i} + \hat{\sigma_1^2}} + 1 \right] > \gamma'$$

Deterministic Signal with unknown param.

- Recall from Section 7.7: Classical linear model

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w}$$

- $\mathbf{H}$ is a N x p known observation matrix
- $\boldsymbol{\theta}$ is a p x 1 vector of unknown signal parameters
- $\mathbf{w}$ is a N x 1 noise vector with PDF $\mathcal{N}(\mathbf{0},\sigma^2\mathbf{I})$

- In contrast to this model, it is assumed that σ^2 is also unknown. Hence, the unknown parameter vector is $[\boldsymbol{\theta}^T\ \sigma^2]^T$
- The use of $\boldsymbol{\theta}$ for the unknown signal parametes only is done to preserve the usual linear model notation.

Deter. Signal with unknown param. (cont.)

- **Theorem 9.1 (GLRT for Classical Linear Model – σ^2 Unknown)**
  - The GLRT for the hypothesis testing problem

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{A}\boldsymbol{\theta} = \mathbf{b}, \sigma^2 > 0 \\ \mathcal{H}_1 &: \mathbf{A}\boldsymbol{\theta} \neq \mathbf{b}, \sigma^2 > 0 \end{aligned}$$

A is a r x p matrix (r ≤ p) of rank r

b is a r x 1 vector of parameters

  - **Aθ** = **b** is a consistent set of linear equations. H_1 is dicided if

$$T(\mathbf{x}) = \frac{N-p}{r}\left(L_G(\mathbf{x})^{\frac{2}{N}} - 1\right)$$

Deter. Signal with unknown param. (cont.)

$$T(\mathbf{x}) = \frac{N-p}{r} \frac{(\mathbf{A}\hat{\boldsymbol{\theta}}_1 - \mathbf{b})^T [\mathbf{A}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T]^{-1}(\mathbf{A}\hat{\boldsymbol{\theta}}_1 - \mathbf{b})}{\mathbf{x}^T(\mathbf{I} - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T)\mathbf{x}} > \gamma'$$

- The exact detection performance (holds for finite data records) is given by

$$\begin{aligned} P_{FA} &= Q_{F_{r,N-p}}(\gamma') \\ P_D &= Q_{F'_{r,N-p}(\lambda)}(\gamma') \end{aligned}$$

- The noncentrality parameter is given by

$$\lambda = \frac{(\mathbf{A}\boldsymbol{\theta}_1 - \mathbf{b})^T [\mathbf{A}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T]^{-1}(\mathbf{A}\boldsymbol{\theta}_1 - \mathbf{b})}{\sigma^2}$$

Example 9.3: Signal with unknown linear parameters in WGN with unknown variance

- The signal is $\mathbf{s} = \mathbf{H}\boldsymbol{\theta}$ and we wish to test if $\mathbf{s} = 0$ versus $\mathbf{s} \neq 0$. If H is full-rank, the equivalent problem is to test $\boldsymbol{\theta} = 0$ versus $\boldsymbol{\theta} \neq 0$.

$$T(\mathbf{x}) = \frac{N-p}{p} \frac{\hat{\boldsymbol{\theta}}_1^T \mathbf{H}^T \mathbf{H} \hat{\boldsymbol{\theta}}_1}{\mathbf{x}^T (\mathbf{I} - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T)\mathbf{x}} > \gamma'$$

$$\begin{cases} \mathbf{A} = \mathbf{I} \\ \mathbf{b} = \mathbf{0} \\ r = p \end{cases}$$

- One intresting interpretation of T(x) follows as

$$\begin{aligned} T(\mathbf{x}) &= \frac{N-p}{p} \frac{[(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}]^T \mathbf{H}^T\mathbf{H} [(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}]}{\mathbf{x}^T (\mathbf{I} - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}} \\ &= \frac{N-p}{p} \frac{\mathbf{x}^T\mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}}{\mathbf{x}^T (\mathbf{I} - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T)\mathbf{x}} \\ &= \frac{N-p}{p} \frac{\mathbf{x}^T\mathbf{P}_H\mathbf{x}}{\mathbf{x}^T\mathbf{P}_H^{\perp}\mathbf{x}} = \frac{N-p}{p} \frac{||\mathbf{P}_H\mathbf{x}||^2}{||\mathbf{P}_H^{\perp}\mathbf{x}||^2} \end{aligned}$$

Example 9.3 (cont.)

- T(x) is an estimated SNR.

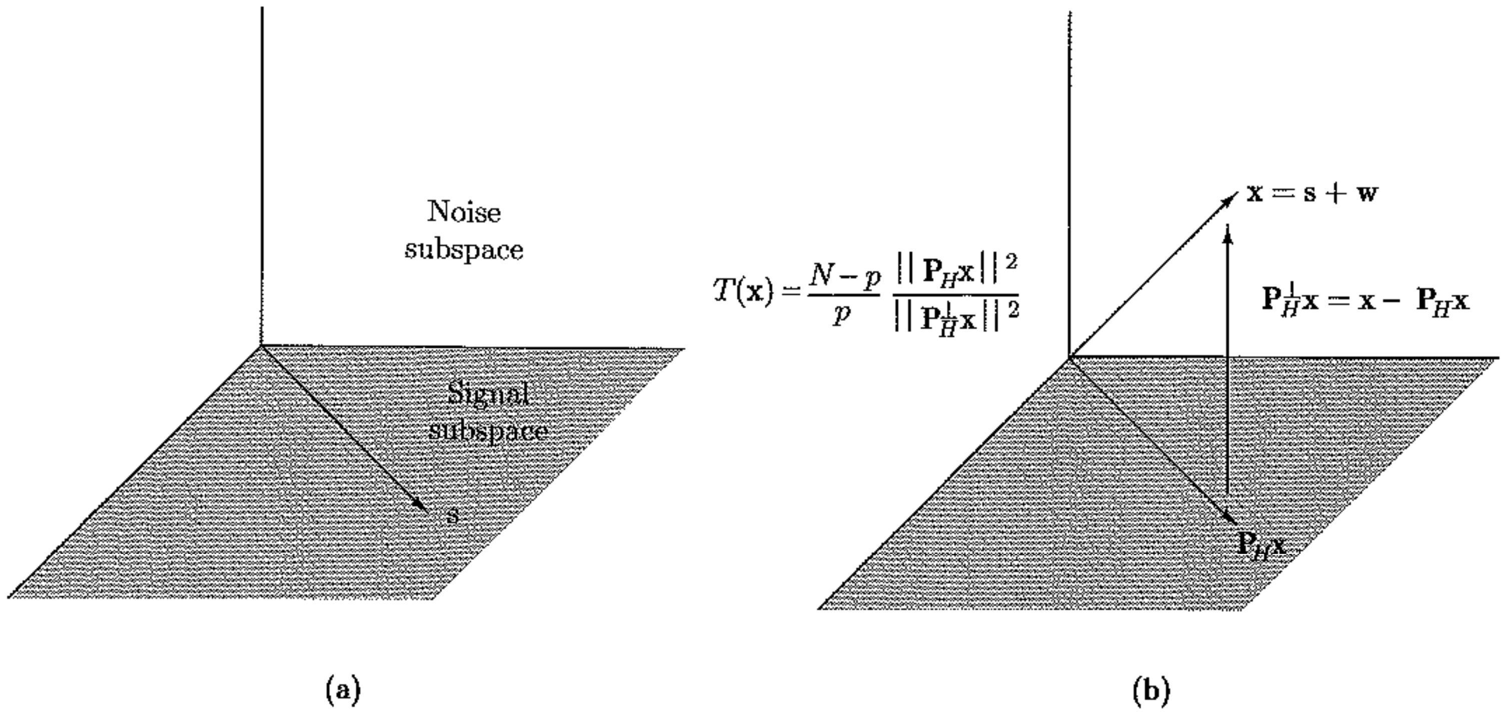


Figure 9.2. Vector space interpretation of detector for linear model.

Colored WSS Gaussian Noise

- The noise is assumed to be WSS.

- A simple model for WSS colored noise is the autoregressive model of order 1

$$w[n] = -a[1]w[n-1] + u[n]$$

$$r_{ww}[k] = \frac{\sigma_u^2}{1 - a^2[1]} (-a[1])^{|k|}$$

$$P_{ww}(f) = \frac{\sigma_u^2}{|1 + a[1]\exp(-j2\pi f)|^2} \qquad |f| \leq \frac{1}{2}$$

It is considered that a[1] and σ_u^2 are unknown

Colored Noise: Deterministic signal

- Before determining the GLRT for this case, we first require the PDF of the noise samples.

- A large data approximation is used.

$$\ln p(\mathbf{w}; a[1], \sigma_u^2) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\ln \sigma_u^2 - \frac{N}{2\sigma_u^2}\int_{-\frac{1}{2}}^{\frac{1}{2}} |A(f)|^2 I_w(f) df$$

where

$$A(f) = 1 + a[1]\exp(-j2\pi f)$$

$$I_w(f) = (1/N)\left|\sum_{n=0}^{N-1} w[n]\exp(-j2\pi f n)\right|^2$$

Colored Noise: Deterministic signal (cont.)

- Based on this approximation, the MLE of a[1] and σ_u^2 are

$$\hat{a}[1] = -\frac{\hat{r}_{ww}[1]}{\hat{r}_{ww}[0]}$$

$$\begin{aligned}\hat{\sigma_u^2} &= \hat{r}_{ww}[0] + \hat{a}[1]\hat{r}_{ww}[1] \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} |\hat{A}(f)|^2 I_w(f)df\end{aligned}$$

where

$$\hat{r}_{ww}[k] = \frac{1}{N}\sum_{n=0}^{N-1-k} w[n]w[n+k] \quad \text{and} \quad \hat{A}(f) = 1 + \hat{a}[1]\exp(-j2\pi f)$$

Thus, we have

$$\ln p(\mathbf{w}; \hat{a}[1], \hat{\sigma_u^2}) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\ln \hat{\sigma_u^2} - \frac{N}{2} \approx p(\mathbf{x}; \hat{a}_0[1], \hat{\sigma_{u_0}^2}, \mathcal{H}_0)$$

Colored Noise: Deterministic signal (cont.)

- The GLRT becomes

$$2 \ln \frac{p(\mathbf{x}; \hat{a}_1[1], \hat{\sigma^2_{u_1}}, \mathcal{H}_1)}{p(\mathbf{x}; \hat{a}_0[1], \hat{\sigma^2_{u_0}}, \mathcal{H}_0)} = N \ln \frac{\hat{\sigma^2_{u_0}}}{\hat{\sigma^2_{u_1}}}$$

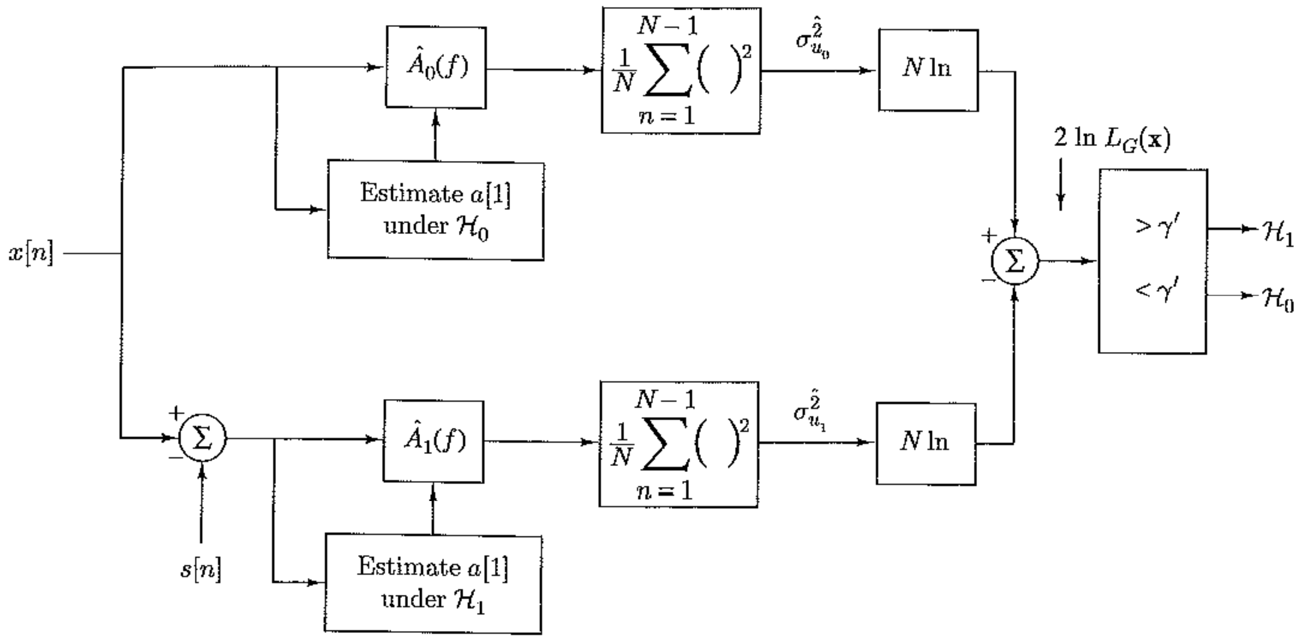
where

$$\begin{aligned} \hat{\sigma^2_{u_0}} &= \int_{-\frac{1}{2}}^{\frac{1}{2}} |\hat{A}_0(f)|^2 I_x(f) df \\ \hat{\sigma^2_{u_1}} &= \int_{-\frac{1}{2}}^{\frac{1}{2}} |\hat{A}_1(f)|^2 I_{x-s}(f) df \end{aligned}$$

and

$$\begin{aligned} \hat{a}_0[1] &= -\frac{\hat{r}_{xx}[1]}{\hat{r}_{xx}[0]} \\ \hat{a}_1[1] &= -\frac{\hat{r}_{x-s,x-s}[1]}{\hat{r}_{x-s,x-s}[0]} \end{aligned}$$

Colored Noise: Deterministic signal (cont.)



$$\hat{A}_0(f) = 1 + \hat{a}_0[1]\exp(-j2\pi f)$$

$$\hat{A}_1(f) = 1 + \hat{a}_1[1]\exp(-j2\pi f)$$

Figure 9.4. GLRT for known deterministic signal in colored autoregressive noise with unknown parameters.

Rao Test for the General Linear Model

- **Theorem 9.2**

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w}$$

- **H** is a N x p known observation matrix with rank p
- **θ** is a p x 1 parameter vector
- **w** is a N x 1 noise vector with PDF $\mathcal{N}(\mathbf{0},\mathbf{C}(\boldsymbol{\theta}_w))$
- $\boldsymbol{\Theta}_w$ is a q x 1 unknown noise parameter

- The Rao test for the hypothesis testing problem

$$\begin{aligned}\mathcal{H}_0 &: \boldsymbol{\theta} = \mathbf{0}, \boldsymbol{\theta}_w \\ \mathcal{H}_1 &: \boldsymbol{\theta} \neq \mathbf{0}, \boldsymbol{\theta}_w\end{aligned}$$

Rao Test (cont.)

- We decides H_1 if

$$T_R(\mathbf{x}) = \mathbf{x}^T \mathbf{C}^{-1}(\hat{\boldsymbol{\theta}}_{w_0})\mathbf{H}(\mathbf{H}^T \mathbf{C}^{-1}(\hat{\boldsymbol{\theta}}_{w_0})\mathbf{H})^{-1}\mathbf{H}^T \mathbf{C}^{-1}(\hat{\boldsymbol{\theta}}_{w_0})\mathbf{x} > \gamma'$$

where $\hat{\boldsymbol{\theta}}_{w_0}$ is the MLE of $\boldsymbol{\theta}_w$ under H0 or the value obatined by maximizing

$$p(\mathbf{x}; \boldsymbol{\theta}_w, \mathcal{H}_0) = \frac{1}{(2\pi)^{\frac{N}{2}} \det^{\frac{1}{2}}(\mathbf{C}(\boldsymbol{\theta}_w))} \exp\left[-\frac{1}{2}\mathbf{x}^T \mathbf{C}^{-1}(\boldsymbol{\theta}_w)\mathbf{x}\right]$$

Rao Test (cont.)

- The asymptotic performance is

$$\begin{aligned} P_{FA} &= Q_{\chi_p^2}(\gamma') \\ P_D &= Q_{\chi_p'^2(\lambda)}(\gamma') \end{aligned}$$

and

$$\lambda = \boldsymbol{\theta}_1^T \mathbf{H}^T \mathbf{C}^{-1}(\boldsymbol{\theta}_w)\mathbf{H}\boldsymbol{\theta}_1$$

- The asymptotic performance is the same if $\boldsymbol{\theta}_w$ is known. When $\boldsymbol{\theta}_w$ is known, the Rao test becomes

$$T_R(\mathbf{x}) = \mathbf{x}^T \mathbf{C}^{-1}(\boldsymbol{\theta}_w)\mathbf{H}(\mathbf{H}^T \mathbf{C}^{-1}(\boldsymbol{\theta}_w)\mathbf{H})^{-1}\mathbf{H}^T \mathbf{C}^{-1}(\boldsymbol{\theta}_w)\mathbf{x} > \gamma'$$

Problems

- 9.2, 9.3, 9.11, 9.19, 9.22