Lecture 10: Signal Detection in Discrete Time

12 February 2016

1 Introduction

In this lecture we apply the theory of hypothesis-testing to detect or discern signals corrupted by noise. We try to build a model which will perform this task of detecting signals embedded in noise. The most common application of this theory is in communication receivers. Some other applications are in the field of radar and sonar receivers, radio astronomy, experimental physics, etc.

2 Models and Detector Structures

The basic physical model we consider is that of an observed continuous-time waveform that consists of one of the two possible signals embedded in noise.

Consider having n samples of the waveform being observed and let the signal be denoted by an n length vector $\underline{Y} = (Y_1, ..., Y_n)^T$. Similarly, let $\underline{N} = (N_1, ..., N_n)^T$ be a vector of noise samples, and $\underline{S}_0 = (S_{01}, ..., S_{0n})^T$ and $\underline{S}_1 = (S_{11}, ..., S_{1n})^T$ be the vectors of samples from the two possible signals.

This problem can be modelled statistically by the following hypothesis for the observation space $(\Gamma,\mathcal{G})=(\mathbb{R}^n, \mathcal{B}^n)$:

$$\begin{aligned} & H_0 : Y_k = N_k + S_{0k}, \quad k = 1, 2, ..., n \\ \text{versus} & \\ & H_1 : Y_k = N_k + S_{1k}, \quad k = 1, 2, ..., n. \end{aligned} \tag{1}$$

Depending on the nature of the signals $\underline{S}_0$ and $\underline{S}_1$, we can have three cases:

1. $\underline{S}_0$ and $\underline{S}_1$ are completely known (i.e., deterministic).
2. $\underline{S}_0$ and $\underline{S}_1$ are partially known except for a set of unknown (possibly random) parameters.

3. $\underline{S}_0$ and $\underline{S}_1$ are completely unknown and thus specified only by their probability distribution.

Assumptions 2.1. $\underline{S}_0 = \underline{0}$ (an all zero vector) and $\underline{S}_1 = \underline{S}$.

Assumptions 2.2. The noise is independent of the signal i.e. at each time instant k, the sample is corrupted by independent noise, and the noise distribution is determined by density p_N on $\mathbb{R}^n$.

With the above discussed framework and assumptions made, the likelihood ratio for (1) can be computed if the statistic of $\underline{S}_j$ for $j = 0, 1$ is known. Given that $\underline{S}_j$=$\underline{s}_j \in \mathbb{R}^{\mathbf{n}}$, the conditional density of $\underline{Y}$ (under H_j) is given by

$$p_{\underline{N}}(\underline{y} - \underline{s}_j), \quad \underline{y} \in \mathbb{R}^{\mathrm{n}}. \tag{2}$$

From (2) we see that the density of $\underline{Y}$ under H_j is given by

$$p_j(\underline{y}) = E\{p_{\underline{N}}(\underline{y} - \underline{S}_j)\}, \quad \underline{y} \in \mathbb{R}^{\mathrm{n}}, \tag{3}$$

where $E\{.\}$ the expectation is with respect to signal $\underline{S}_j$. The general expression of the likelihood ratio for (1) is given by

$$L(\underline{y}) = \frac{E\{p_{\underline{N}}(\underline{y} - \underline{S}_1)\}}{E\{p_{\underline{N}}(\underline{y} - \underline{S}_0)\}}, \quad \underline{y} \in \mathbb{R}^{\mathrm{n}}. \tag{4}$$

2.1 Case I: Detection of Deterministic Signals in Independent Noise

As discussed earlier, here the two signals $\underline{S}_0$ and $\underline{S}_1$ are completely known or deterministic. In the field of communication, this is known as the *coherent detection* problem. Here, $L(\underline{y})$ of (4) becomes

$$\begin{aligned} L(\underline{y}) &= \frac{p_{\underline{N}}(\underline{y} - \underline{s}_1)}{p_{\underline{N}}(\underline{y} - \underline{s}_0)}, \\ &= \frac{p_{\underline{N}}(y_1 - s_{11}, y_1 - s_{12}, ..., y_1 - s_{1n})}{p_{\underline{N}}(y_1 - s_{01}, y_1 - s_{12}, ..., y_1 - s_{0n})}, \end{aligned} \tag{5}$$

since the noise samples $N_1, N_2, ..., N_n$ are statistically independent (by assumption). So we have

$$p_{\underline{N}}(\underline{y}) = \prod_{k=1}^{n} p_{N_k}(y_k), \tag{6}$$

where p_{N_k} is the marginal density of N_k. So, $L(\underline{y})$ of (5) becomes

$$L_(\underline{y}) = \prod_{k=1}^{n} \frac{p_{N_k}(y_k - s_{1k})}{p_{N_k}(y_k - s_{0k})}. \tag{7}$$

Comparing this $L(\underline{y})$ to a threshold τ gives us the decision rule

$$\tilde{\delta}_o(\underline{y}) = \begin{cases} 1 & & < & \\ \gamma & \text{if } L(\underline{y}) & = & \tau. \\ 0 & & > & \end{cases} \tag{8}$$

Example 2.3 (Coherent Detection in i.i.d Gaussian Noise). Suppose that the noise samples $N_1, ..., N_n$ are independent and identically distributed (i.i.d.) with marginal distribution $\mathcal{N}(0, \sigma^2)$. Without any loss of generality, assuming $\underline{s}_0 = \underline{0}$ (all zero vector) and $\underline{s}_1 = \underline{s}$, $L(\underline{y})$ of (7) can be written as

$$L(\underline{y}) = \prod_{k=1}^{n} \exp\left(-\frac{(y_k - s_k)^2}{2\sigma^2} + \frac{(y_k)^2}{2\sigma^2}\right). \tag{9}$$

Taking log on both sides, log $L(\underline{y})$ can be expressed as

$$\begin{aligned} \log L(\underline{y}) &= \sum_{k=1}^{n} -\frac{(y_k - s_k)^2}{2\sigma^2} + \frac{(y_k)^2}{2\sigma^2}, \\ &= \sum_{k=1}^{n} \frac{-(y_k^2 + s_k^2 - 2y_k s_k) + y_k^2}{2\sigma^2}, \\ &= \sum_{k=1}^{n} \frac{2y_k s_k - s_k^2}{2\sigma^2}, \\ &= \frac{1}{\sigma^2} \sum_{k=1}^{n} s_k (y_k - s_k/2). \end{aligned} \tag{10}$$

Thus the decision rule becomes

$$\tilde{\delta}_o(\underline{y}) = \begin{cases} 1 & & < & \\ \gamma & \text{if } \sum_{k=1}^{n} s_k (y_k - s_k/2) & = & \sigma^2 \log \tau, \\ 0 & & > & \end{cases} \tag{11}$$

or equivalently

$$\tilde{\delta}_o(\underline{y}) = \begin{cases} 1 & & < & \\ \gamma & \text{if } \sum_{k=1}^{n} s_k y_k & = & \tau', \\ 0 & & > & \end{cases} \tag{12}$$

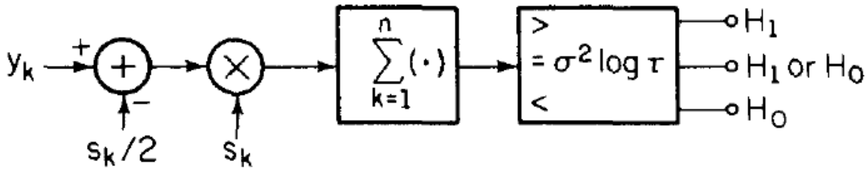

Figure 1: Optimum detector for coherent signals i.i.d Gaussian noise

(*Source: H. Vincent Poor, An Introduction to Signal Detection and Estimation(Second Edition), Figure III.B.2(a)*)

where $\tau' \triangleq \sigma^2 \log \tau + \frac{1}{2}\sum_{k=1}^{n} s_k^2$. This detector structure is depicted in Fig. 1.

This system centers the observation by subtracting $s_k/2$ from each y_k. It correlates the centered data with the known signal and compares the output of this correlation with a threshold. It can be viewed as a system that inputs the observation sequence $(y_1, y_2 ...., y_n)$ to a linear digital filter and then samples the output at time n for comparison with a threshold. Such a structure is known as a *matched filter.*

Example 2.4 (Coherent Detection in i.i.d Laplacian Noise). Suppose, as in Example 2.3, that the noise samples $N_1, ..., N_n$ are i.i.d but with Laplacian marginal probability density

$$p_{N_k}(y_k) = \frac{\alpha}{2} e^{-\alpha|y_k|}, \quad y_k \in \mathbb{R}, \tag{13}$$

where $\alpha > 0$ is a scale parameter of the density. This model is sometimes used to represent the behavior of impulsive or burst noise in communication receivers.

The function $\log L_k(y_k)$ for (13) is given by $\log L_k(y_k) = \alpha(|y_k| - |y_k - s_k|)$, which can be written as

$$\log L_k(y_k) = \begin{cases} -\alpha|s_k| & \text{if sgn } (s_k)y_k \leq 0 \\ \alpha|2y_k - s_k| & \text{if } 0 < \text{sgn } (s_k)y_k < |s_k| \\ +\alpha|s_k| & \text{if sgn } (s_k)y_k \geq |s_k|, \end{cases} \tag{14}$$

where *sgn* denotes the signum function

$$\text{sgn } (x) = \begin{cases} +1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases} \tag{15}$$

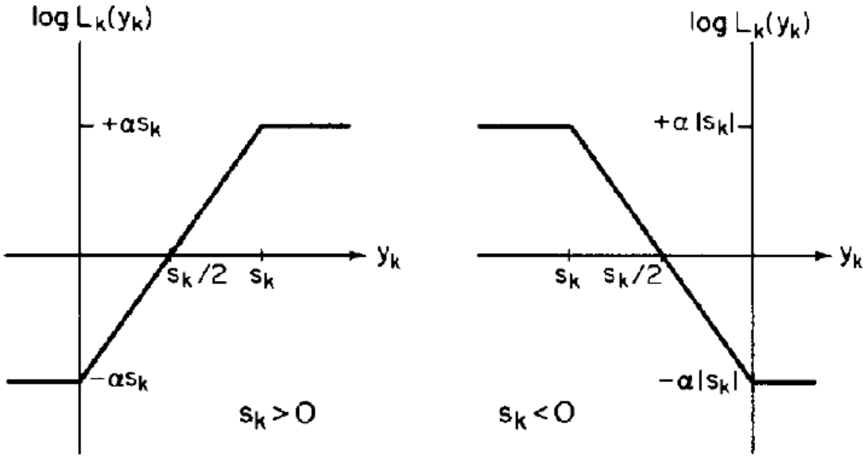

Figure 2: Per-sample log-likelihood ratio for coherent detection in Laplacian Noise

(*Source: H. Vincent Poor, An Introduction to Signal Detection and Estimation(Second Edition), Figure III.B.3*)

This function $\log L_k(y_k)$ for both cases ($s_k > 0$ *and* $s_k < 0$) is depicted in Fig. 2. By inspection of these figures the decision rule can be written as

$$\tilde{\delta}_o(\underline{y}) = \begin{cases} 1, & & > \\ \gamma, & \text{if } \sum_{k=1}^{n} \text{sgn } (s_k) l_k(y_k - s_k/2) & = \tau, \\ 0, & & < \end{cases} \tag{16}$$

where the function l_k is given by

$$l_k(x) = \begin{cases} -|s_k|/2, & \text{if } x \leq -|s_k|/2, \\ x, & \text{if } -|s_k|/2 < x < |s_k|/2, \\ +|s_k|/2, & \text{if } x \geq +|s_k|/2. \end{cases} \tag{17}$$

Such a function is known as a *soft limiter/amplifier limiter*. The structure for such a model is depicted in Fig. 3.

This detector also centers the observation by subtracting $s_k/2$ from each y_k. Then it soft limits the centered data and then correlates these soft limited observations with the sequence of signal signs. The effect of soft limiting is to reduce the effect of large observations on the sum, thus making the system more tolerant to large noise values.

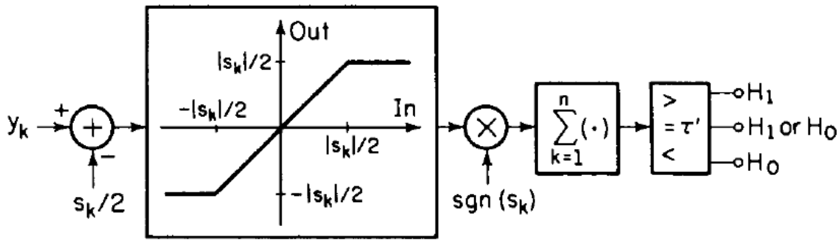

Figure 3: Optimum detector for coherent signals in Laplacian Noise

(*Source: H. Vincent Poor, An Introduction to Signal Detection and Estimation(Second Edition), Figure III.B.4*)

Example 2.5 (Locally Optimum Detection of Coherent Signals in i.i.d. Noise). Often in many detection problems the form of the received signals is known but not its amplitude. To model such a problem we consider the composite hypothesis-testing problem given by:

$$\begin{array}{c} H_0 : Y_k = N_k, \quad k = 1, 2, ..., n \\ \text{versus} \\ H_1 : Y_k = N_k + \theta_{S_k}, \quad k = 1, 2, ..., n, \quad \theta > 0, \end{array} \tag{18}$$

where $\underline{s} = (s_1, ..., s_n)^T$ is a known signal, $\underline{N} = (N_1, ..., N_n)^T$ is a continuous random vector with i.i.d. components and marginal probability density functions p_{N_k}, and θ is an unknown signal-strength parameter, i.e. signal s_k should have been scaled with unknown amplitude θ, where the distribution is

$$\begin{aligned} \Lambda_0 &= \{0\}, \quad \theta = 0 \\ \Lambda_1 &= (0, \infty), \quad \theta > 0. \end{aligned}$$

For any particular (given) θ, the $L_\theta(\underline{y})$ of (18) is given by

$$L_\theta(\underline{y}) = \prod_{k=1}^{n} \frac{p_{N_k}(y_k - \theta s_k)}{p_{N_k}(y_k)}. \tag{19}$$

The critical region for any θ, $\Gamma_\theta = \{\underline{y} \in \mathbb{R}^n | L_\theta(\underline{y}) > \tau\}$ will generally depend on θ. Hence a *uniformly most powerful* (UMP) test may not exist. Instead, we can try to find a *locally most powerful* (LMP) test. It is in some sense optimal when θ is very close to 0.

Note 1. Given $\alpha \in$ (0,1) the LMP test carries out

$$\max_\delta \left(P'_D(\delta, \theta_o) = \frac{\partial}{\partial \theta} P_D(\delta, \theta) \bigg|_{\theta = \theta_o} \right),$$

such that $P_F(\delta) \leq \alpha$.

Thus, the form of LMP will be

$$\tilde{\delta}_{lo}(y) = \begin{cases} 1 & > \\ \gamma \quad \text{if } \frac{\partial}{\partial\theta} P_\theta(y) \left|_{\theta=\theta_o} \right. & = \eta . P_{\theta_o}(y). \\ 0 & < \end{cases} \tag{20}$$

Upon differentiation of (19), we have

$$\begin{aligned}
\frac{\partial}{\partial\theta} L_\theta(\underline{y}) \left|_{\theta=0}\right. &= \frac{\partial}{\partial\theta} \left(\prod_{k=1}^{n} \frac{p_{N_1}(y_k - \theta_{S_k})}{p_{N_1}(y_k)} \right) \Bigg|_{\theta=0} \\
&= \left(\prod_{k=1}^{n} \frac{p_{N_1}(y_k - \theta_{S_k})}{p_{N_1}(y_k)} \right) \left(\sum_{k=1}^{n} \frac{\frac{\partial}{\partial\theta} p_{N_1}(y_k - \theta_{S_k})}{p_{N_1}(y_k - \theta_{S_k})} \right) \Bigg|_{\theta=0} \\
&= \sum_{k=1}^{n} \frac{-s_k p'_{N_1}(y_k - \theta_{S_k})}{p_{N_1}(y_k - \theta_{S_k})} \Bigg|_{\theta=0} \\
&= \sum_{k=1}^{n} s_k \left(\frac{-p'_{N_1}(y_k)}{p_{N_1}(y_k)} \right) \\
&= \sum_{k=1}^{n} s_k g_{lo}(y_k),
\end{aligned} \tag{21}$$

where $g_{lo}(x) \triangleq -\, p'_{N_1}(x)/p_{N_1}(x)$, and where $p'_{N_1}(x) = dp_{N_1}(x)/dx$. This structure

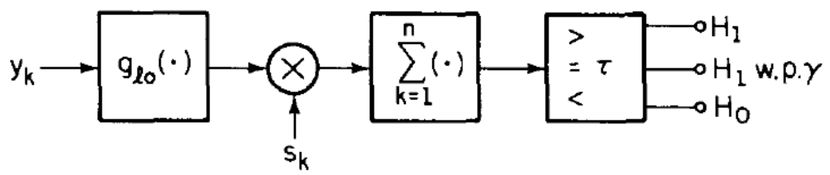

Figure 4: Locally optimum detector structure for coherent signals in i.i.d noise

(*Source: H. Vincent Poor, An Introduction to Signal Detection and Estimation(Second Edition), Figure III.B.5*)

is depicted in Fig. 4. It consists of the memoryless non-linearity g_{lo} followed by a correlator, a combination known as a *nonlinear correlator*.

Like the likelihood ratio, the locally optimum nonlinearity g_{lo}, shapes the observations to reduce the detrimental effects of the noise as much as is possible. For example, with $\mathcal{N}(0, \sigma^2)$ noise, we have $g_{lo}(x) = x/\sigma^2$, so that Fig. 4 is simply the

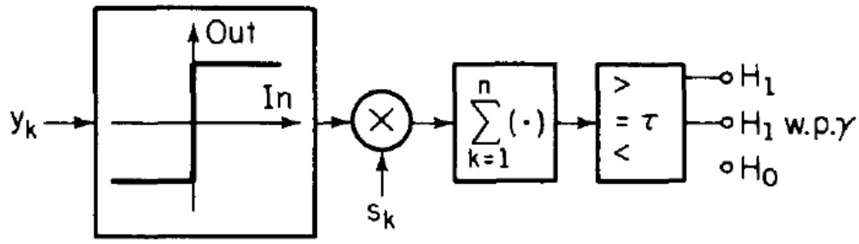

Figure 5: Locally optimum detector for Laplacian noise

(*Source: H. Vincent Poor, An Introduction to Signal Detection and Estimation(Second Edition), Figure III.B.6*)

correlation detector of Fig.1. This must be so; since this detector is UMP, it is also LMP.

For Laplacian noise with density (13) we have $g_{lo}(x) = \alpha$ sgn (x). The locally optimum detector correlates the signal with the sequences of signs of the observations as depicted in Fig. 5. The function $g_{lo}(x)$ in this case is known as a *hard limiter*.