EECS 555: Digital Communication Theory

Wayne E. Stark

Contents

Chapter 1

Introduction

The goal of communication systems is to reliably transmit information from one location to another. This can be done in various ways and depends on certain resources. These resources include the energy, and the bandwidth of the channel. However, the channel impairments and limitations on complexity of the design limit the capabilities to reliably transmit information. Below we discuss various parameters and the effect on reliable communications.

- **Power or Energy:** Clearly the more power available the more reliable communication is possible. However, the goal is to achieve reliable communication with the minimum required transmission power. For cell phones this maximizes the talk time.
- **Data Rate:** The goal is large data rates. However, for a fixed amount of power as the data rate increases the energy transmitted per bit will decrease because of decreased transmission time for each bit. In addition, if the data rate increases then the amount of intersymbol interference will increase. A wireless channel typically has an impulse response with some delay spread. That is, the received signal is delayed by different amounts on different paths. The signal corresponding to a particular bit received with the longest delay with interfere with the signal corresponding to a different bit with the shortest delay. The larger the number bits that are interfered with the more difficult it is to correct for this interference.
- **Bandwidth:** This is the amount of frequency spectrum available for use. Generally the FCC allocates spectrum and provides some type of mask for which the radios emissions must fall within. The larger the bandwidth the more indendent fades accross frequencies and thus better averaging is possible.
- **Probability of Error:** Usually the probability of bit error is the performance measure of interest although probability of packet error is of interest in some systems. Clearly data communications requires low error probabilities than voice, but also allows more delay.
- **Delay Spread (Coherence Bandwidth):** The delay spread of a channel measures the differential delay between the longest significant path and the shortest significant path in a channel. The delay spread is inversely related to the coherence bandwidth which indicates the minimun frequency separation such that the response at the two different frequencies is independent.
- **Coherence Time (Doppler Spread):** This is related to the vehicular speed. The correlation time measures how fast the channel is changing. If the channel changes quickly it is hard to estimate the channel response. However a quickly changing channel also ensures that a deep fade does not last too long. The Doppler spread is the frequency characteristics of the channel impulse response and it is inversely related to the correlation time.
- **Delay Requirment:** Larger delay requirements allow for larger number of fades to be averaged out.
- **Complexity:** More complexity usually implies better performance. The trick is to get the best for less.

The overall design of a communication system depends on the relative importance of different parameters (energy, delay, etc.). The goal of this book is to understand the tradeoff possible in designing a communication system between these parameters.

1. Communication System Coat of Arms

There are many different functions in a digital communication system. These are represented in the block diagram shown below.

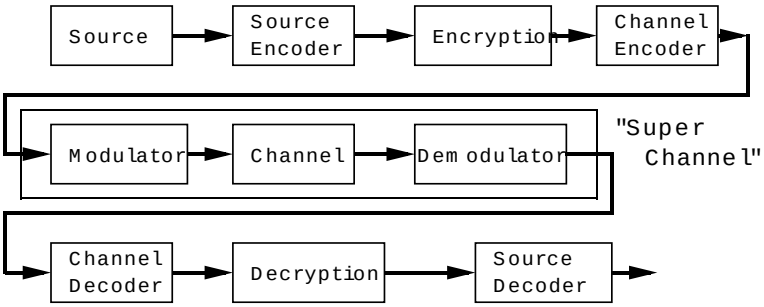


Figure 1.1: Block Diagram of a Digital Communication System

- **Source Encoder:** Removes redundancy from the source data such that the output of the source encoder is a sequence of symbols from a finite alphabet. If the source produces symbols from an infinite alphabet than some distortion must be incurred in representing the source with a finite alphabet. If the rate at which the source produces symbols is below the ”entropy” of the source than distortion must be incurred.

- **Encryption Device** Transforms input sequence $\{W_k\}$ into an output sequence $\{Z_n\}$ such that knowledge of $\{Z_n\}$ alone (without a key) makes calculation of $\{W_l\}$ extremely difficult (many years of CPU time on a fast computer).

- **Channel Encoder:** Introduces redundancy into data such that if there are some errors made over the channel they can be corrected.

  Note: The source encoder removes *unstructured* redundancy from the source data and may cause distortion or errors in a *controlled* fashion. The channel encoder adds redundancy in a structured fashion so that the channel decoder can correct some errors caused by the channel.

- **Modulator:** Maps a finite number of messages into a set of distinguishable signals so that at the channel output it is possible to determine which signal in the set was transmitted.

- **Channel:** Medium by which signal propagates from transmitter to receiver

  Examples of communication channels:

  Noiseless channel (very good, but not interesting).

  Additive white Gaussian noise channel (classical, for example the deep space channel is essential an AWGN channel).

  Intersymbol interference channel (e.g. the telephone channel)

  Fading channel (mobile communication system when transmitters are behind buildings, Satellite systems when there is rain on the earth).

  Multiple-access interference (when several users access the same frequency at the same time).

  Hostile interference (jamming signals).

  Semiconductor memories (RAM’s, errors due to alpha particle decay in packaging).

  Magnetic and Optical disks (Compact digital disks for audio and for read only memories, errors due to scratches and dust).

- **Demodulator:** Processes the channel output and produces an estimate of the message that caused the output.

- **Channel Decoder:** Reverses the operation of the channel encoder in the absence of any channel noise. When the channel causes some errors to be made in the estimates of the transmitted messages the decoder corrects these errors.

- **Decryption Device:** With the aid of a secret key reverses the operation of the encryption device. With private key cryptography the key determines the method of encryption which is easily invertible to obtain the decryption. With public key cryptography there is a key which is made public. This key allows anyone to encrypt a message. However, even knowing this key it is not possible to reverse this operation (at least not easily) and recover the message from the encrypted message. There are some special properties of the encryption algorithm known only to the decryption device which makes this operation easy. This is known as a trap door. Since the encryption key need not be kept secret for the message to be kept secret this is called public key cryptography.

- **Source Decoder:** Reverse the operation of the source encoder to determine the most probable sequence that could have caused the output.

  Often the modulator-channel-demodulator are thought of as a *super channel* with a finite number of inputs and a finite or infinite number of outputs.

More than 50 years ago Claude Shannon (U of M EE/Math graduate) determined the tradeoff between data rate, bandwidth, signal power and noise power for reliable communications for an additive white Gaussian noise channel. Let W be the bandwidth (in Hz), R be the data rate (in bits per second), P be the *received* signal power (in watts) and $N_0/2$ the noise power spectral density (in watts/Hz) then reliable communication is possible provided

$$R < W \log_2(1 + \frac{P}{N_0 W}).$$

Let E_b be the energy transmitted per bit of information. Then

$$E_b = P/R \quad \text{or} \quad P = E_b R.$$

Using this relation we can express the capacity formula as

$$R/W < \log_2(1 + \frac{E_b}{N_0}\frac{R}{W}).$$

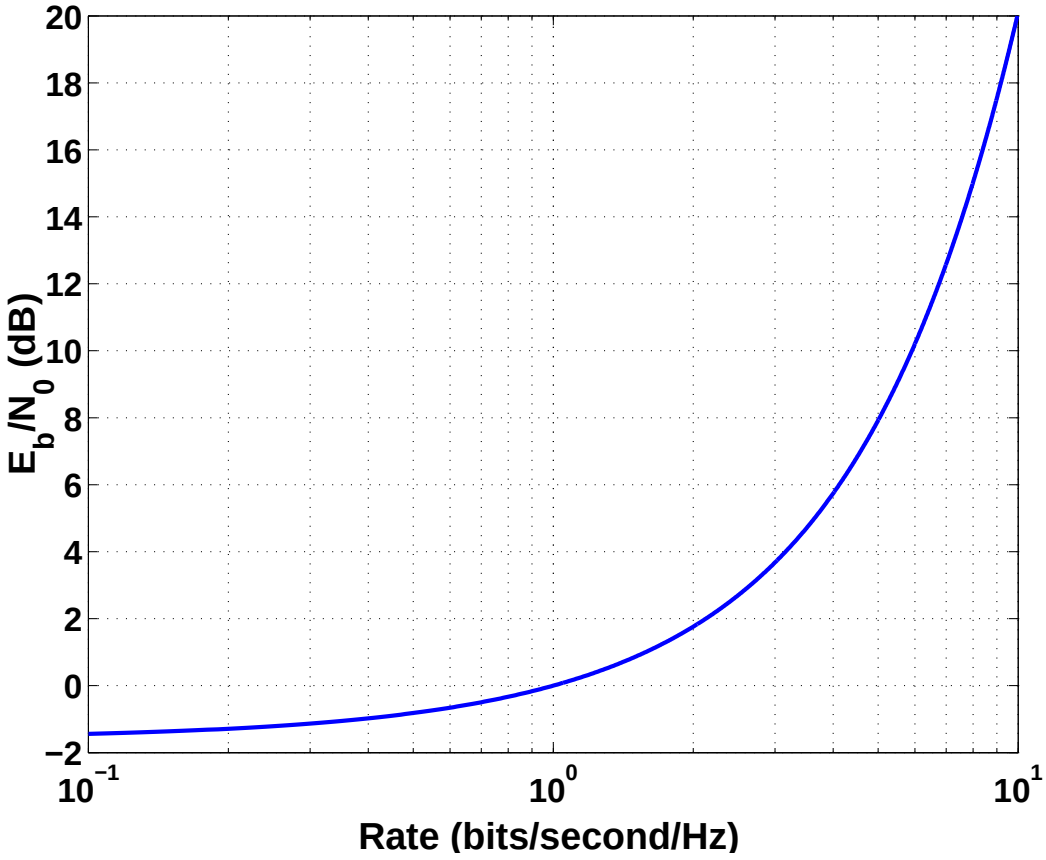
Inverting this we obtain

$$E_b/N_0 > \frac{2^{R/W} - 1}{R/W}.$$

The interpretation is that reliable communication is possible with *bandwidth efficiency* R/W provided that the *signal-to-noise ratio* E_b/N_0 is larger than the right hand side of the above equation. Usually energy or power ratios are expressed in dB's. The conversion is

$$E_b/N_0(dB) = 10\log_{10}(E_b/N_0).$$

The capacity formula only provides a tradeoff between energy efficiency and bandwidth efficiency. Complexity is essentially infinite, as is delay. The model of the channel is rather benign in that no signal fading is assumed to occur.


20
18
16
14
12
10
8
6
4
2
0
−2
E_b/N_0 (dB)
10^{-1}
10^{0}
10^{1}
Rate (bits/second/Hz)

Chapter 2

Optimum Receiver Principles

In this chapter we discuss optimum receiver principles for a digital communication system. We begin by considering the optimum decision rule for deciding which of two hypothesis is true based on an observation. The observation is either a single random variable or a finite number of random variables. The decision rule depends only on the conditional density functions of the observations given each of the two hypothesis. The rule that minimizes the average probability of error is derived. This is then extended to random processes received over a Gaussian (but not necessarily white) noise channel and the M-ary detection where there are M (finite) hypotheses.

1. Detection Theory

In this section we consider the problem of optimally deciding which of M hypotheses are true based on an observation. The observation consists of either a single random variable or a finite length random vector. The distribution (or density) of these random vectors depends upon which hypothesis is true. Thus we assume some conditional distribution or density function for the random variable(s) given the hypothesis. The criteria for the optimization problem for a digital communication system is usually the average probability of error where the average is taken with respect to the probability distribution of the different hypotheses. We begin with a more general cost function or optimization criteria but specialize to the error probability criteria.

1. Binary Detection

The set up for the problem is as follows.

$$\begin{array}{llll} H_0: & X & \text{has density} & p_0(x) \\ H_1: & X & \text{has density} & p_1(x) \end{array}$$

The problem we would like to solve is the following. Upon observing that $X(\omega) = x$ how do we decide which hypothesis is true: H_0 or H_1?

Ex 1. p_0 is Gaussian mean -1 variance 1.

p_1 is Gaussian mean $+1$ variance 1.

Ex 2. p_0 is Gaussian mean 0 variance 1.

p_1 is Gaussian mean 0 variance 10.

We choose H_0 or H_1 depending on x to minimize a cost function. Let c_{ij} = cost of deciding H_i when H_j is true. Assume cost of making wrong decision is greater than cost of making correct decision, i.e.

$$c_{10} > c_{00} \text{ and } c_{01} > c_{11} \ .$$

Let

$$\begin{aligned}\pi_0 &= \text{probability that } H_0 \text{ is true.}\\ \pi_1 &= \text{probability that } H_1 \text{ is true.}\end{aligned}$$

Let $C =$ cost of decision. Then

$$\begin{aligned}E[C] &= \sum_{i=0}^{1}\sum_{j=0}^{1} c_{ij} P\{\text{decide } H_i | H_j \text{ true}\} P\{H_j\}\\ &= c_{00}\, P\{\text{decide } H_0|H_0\}P\{H_0\}\\ &\quad + c_{01}\, P\{\text{decide } H_0|H_1\}P\{H_1\}\\ &\quad + c_{10}\, P\{\text{decide } H_1|H_0\}P\{H_0\}\\ &\quad + c_{11}\, P\{\text{decide } H_1|H_1\}P\{H_1\} \quad .\end{aligned}$$

Let P_F denote the probability of deciding H_1 when actually H_0 is true. This is usually called the probability of false alarm. Let P_D be the probability of deciding H_1 is true when H_1 is true. This is called the probability of detection. Let P_M be the probability of deciding H_0 is true when H_1 is true. This is called the probability of miss.

$$\begin{aligned}P_F &= P\{\text{decide } H_1|H_0\} = 1 - P\{\text{decide } H_0|H_0\}\\ P_D &= P\{\text{decide } H_1|H_1\} = 1 - P\{\text{decide } H_0|H_1\} \ .\end{aligned}$$

The average cost can be written in terms of the probability of false alarm and the probability of detection as follows.

$$\begin{aligned}E[C] &= c_{00}[1-P_F]\pi_0 + c_{01}[1-P_D]\pi_1 + c_{10}[P_F]\pi_0 + c_{11}P_D\pi_1\\ &= c_{00}\pi_0 + c_{01}\pi_1 + (c_{10}-c_{00})\pi_0\, P_F + (c_{01}-c_{11})\, P_D\pi_1 \quad .\end{aligned}$$

The term $c_{00}\pi_0 + c_{01}\pi_0$ represents a fixed cost that does not depend on what the decision made is nor does it depend on what the observation X is. Each of the last two terms are nonnegative.

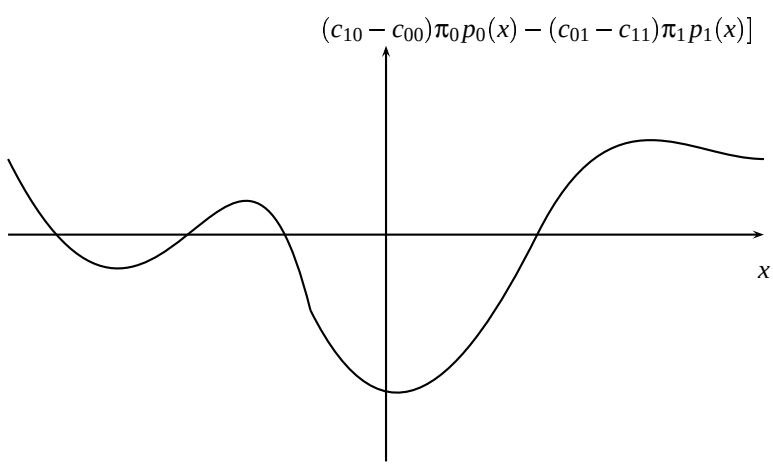
Let $R_0 =$ the region of real line such that if $X(\omega) = x \in R_0$ then the decision H_0 is made. Similarly let $R_1 =$ the region of real line such that if $X(\omega) = x \in R_1$ then the decision H_1 is made. We now write the false alarm and detection probabilities in terms of these regions.

$$P_F = \int_{R_1} p_0(x)dx, P_D = \int_{R_1} p_1(x)dx.$$

The average cost can now be written as

$$E[C] = c_{00}\pi_0 + c_{01}\pi_1 + \int_{R_1} [(c_{10}-c_{00})\pi_0 p_0(x) - (c_{01}-c_{11})\pi_1 p_1(x)]dx.$$

We wish to *minimize* the average cost. The expected cost as seen above is a constant $c_{00}\pi_0 + c_0 1\pi_1$ (independent of the decision region) plus an integral over the region R_1 of a function depending on the densities of the observation given each hypothesis.



We can minimize the average cost by choosing the decision region R_1 to be that region for which the integrand is negative. Thus the optimal decision rule is to decide H_1 if $X = x$ and

$$(c_{10} - c_{00})\pi_0 p_0(x) - (c_{01} - c_{11})\pi_1 p_1(x) < 0.$$

Let

$$\Lambda(x) = \frac{p_1(x)}{p_0(x)}.$$

The optimal decision rule can be rewritten as

$$\begin{aligned} \text{Decide } H_1 \quad &\text{if} \quad \Lambda(X) > \frac{(c_{10} - c_{00})\pi_0}{(c_{01} - c_{11})\pi_1} \\ \text{Decide } H_0 \quad &\text{if} \quad \Lambda(X) < \frac{(c_{10} - c_{00})\pi_1}{(c_{01} - c_{11})\pi_0}. \end{aligned}$$

Equivalently

$$\Lambda(X) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{(c_{01} - c_{11})\pi_0}{(c_{10} - c_{00})\pi_1}.$$

$\Lambda(X)$ is called the likelihood ratio. This is Bayes solution if π_0, π_1 are known to the person designing the test.

If we assign costs so that the performance measure is the average error probability ($c_{0,1} = c_{1,0} = 1,\ c_{0,0} = c_{1,1} = 0$) then the likelihood ratio test has the form

$$\begin{aligned} \Lambda(X) = \frac{p_1(X)}{p_0(X)} &\underset{H_0}{\overset{H_1}{\gtrless}} \frac{\pi_0}{\pi_1} \\ p_1(X)\pi_1 &\underset{H_0}{\overset{H_1}{\gtrless}} p_0(X)\pi_0 \end{aligned}$$

Decision rule is choose i so that $p_i(x)\pi_i > p_j(x)\pi_j\ \forall j \neq i$. Equivalently the optimal decision rule can be described as choose i so that $p(i|x) > p(j|x)\ \ \forall j \neq i$ where $p(i|x)$ is the aposteriori probability of hypothesis i given the observation x. The decision rule is thus to choose the hypothesis that maximizes the aposteriori probability. This is called the maximum aposteriori probability rule or MAP rule.

Example 1:

$$\begin{aligned} p_0(x) &= ae^{-ax}. \\ p_1(x) &= be^{-bx}. \end{aligned}$$

where $a > b$. The likelihood ratio for this test is

$$\Lambda(x) = \frac{p_1(x)}{p_0(x)} = \frac{be^{-bx}}{ae^{-ax}} = \frac{b}{a}\exp\{x(a-b)\} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{(c_{10} - c_{00})\pi_0}{(c_{01} - c_{11})\pi_1} \triangleq \gamma$$

$$\begin{aligned} \frac{b}{a}\, e^{x(a-b)} &\underset{H_0}{\overset{H_1}{\gtrless}} \gamma \\ e^{x(a-b)} &\underset{H_0}{\overset{H_1}{\gtrless}} \frac{a\gamma}{b} \\ x(a-b) &\underset{H_0}{\overset{H_1}{\gtrless}} \ln\frac{a\gamma}{b} \\ x &\underset{H_0}{\overset{H_1}{\gtrless}} \frac{\ln(a\gamma/b)}{(a-b)}. \end{aligned}$$

The probability of false alarm and detection for the optimal decision rule can be determined as follows.

$$\begin{aligned} P_F &= P\{\Lambda(X) > \gamma | H_0\} = P\{X > \frac{\ln(a\gamma/b)}{(a-b)} | H_0\} \\ &= \int_{\frac{\ln(a\gamma/b)}{(a-b)}}^{\infty} a e^{-ax} dx = -e^{-ax}\Big|_{\frac{\ln(a\gamma/b)}{(a-b)}}^{\infty} \\ &= e^{-\frac{a\ln(a\gamma/b)}{(a-b)}}. \end{aligned}$$

$$\begin{aligned} P_D &= P\{\Lambda(X) > \gamma | H_1\} = P\{X > \frac{\ln(a\gamma/b)}{a-b} | H_1\} \\ & \int_{\frac{\ln(a\gamma/b)}{a-b}}^{\infty} b e^{-bx} = e^{-bx}\Big|_{\frac{\ln(a\gamma/b)}{(a-b)}}^{\infty} \\ &= \exp\left\{-\frac{b\ln(a\gamma/b)}{(a-b)}\right\}. \end{aligned}$$

$$E[C] = [c_{00} + (c_{10} - c_{00})P_F]\pi_0 + [c_{11} - (c_{01} - c_{11})P_D]\pi_1.$$

For error probability criterion $c_{00} = c_{11} = 0 \;\; c_{01} = c_{10} = 1$.
Example 2:

$$\begin{aligned} H_0 &: X = (00000000\ldots0) \oplus \underline{N} \\ H_1 &: X = (11111111\ldots1) \oplus \underline{N} \end{aligned}$$

where $N = (N_1, \ldots N_L)$ is a random noise vector with each component being 0 or 1 with $P\{N_1 = 0\} = 1 - \rho$, $P\{N_1 = 1\} = \rho$. Also $N_1, \ldots, N_L$ is an i.i.d. sequence. The likelihood ratio for the observation X can be determined as follows.

$$\begin{aligned} P\{X = x | H_0\} &= \rho^j (1-\rho)^{L-j} \quad w_H(x) = j \\ P\{X = x | H_1\} &= \rho^{L-j} (1-\rho)^j \quad w_H(x) = j \\ \Lambda(j) = \frac{P\{X \text{ has } j \text{ ones } | H_1\}}{P\{X \text{ has } j \text{ ones } | H_0\}} &= \frac{\binom{L}{L-j}\rho^{L-j}(1-\rho)^j}{\binom{L}{j}\rho^j(1-\rho)^{L-j}} \\ \left(\frac{\rho}{1-\rho}\right)^{L-2j} &\underset{H_0}{\overset{H_1}{\gtrless}} 1 \\ (L-2j)\log[\rho/(1-\rho)] &\underset{H_0}{\overset{H_1}{\gtrless}} 0. \end{aligned}$$

Assume $\rho < 1/2 \Rightarrow \rho/(1-\rho) < 1 \Rightarrow \log(\rho/(1-\rho)) < 0$. Thus the likelihood ratio test becomes

$$j \underset{H_0}{\overset{H_1}{\gtrless}} \frac{L}{2}.$$

If $j = L/2$ then flip a coin. (This is obvious rule!)
Example 3:

$$\begin{aligned} H_0 &: X_i = n_i && i = 0, 1, 2, \ldots, N \\ H_1 &: X_i = s_i + n_i && i = 0, 1, 2, \ldots, L \\ & X_i = n_i && i = L+1, \ldots, N \end{aligned}$$

where $n_i,\ 0 \le i \le N$ is sequence of i.i.d. Gaussian random variables mean 0 variance σ^2.

$$\begin{aligned}
p_0(\underline{x}) = \prod_{i=0}^{N} p_0(x_i) &= \prod_{i=0}^{N} \frac{1}{\sqrt{2\pi}\sigma}\, e^{-(x_i^2/2\sigma^2)} \\
&= \frac{1}{(\sqrt{2\pi}\sigma)^N}\, \exp\left\{ -\sum_{i=0}^{N} x_i^2/2\sigma^2 \right\} \\
p_1(\underline{x}) = \prod_{i=0}^{N} p_1(x_i) &= \prod_{i=0}^{L} \frac{1}{\sqrt{2\pi}\sigma}\, e^{-((x_i-s_i)^2/2\sigma^2)} \prod_{i=L+1}^{N} \frac{1}{\sqrt{2\pi}\sigma}\, e^{-x_i^2/2\sigma^2} \\
&= \frac{1}{(\sqrt{2\pi}\sigma)^N}\, \exp\left\{ -\frac{1}{2\sigma^2} \left[\sum_{i=0}^{L} (x_i - s_i)^2 + \sum_{i=L+1}^{N} x_i^2 \right] \right\}
\end{aligned}$$

$$\begin{aligned}
\frac{p_1(\underline{x})}{p_0(\underline{x})} &= \exp\left\{ -\frac{1}{2\sigma^2} \left[\sum_{i=0}^{L} (x_i - s_i)^2 - x_i^2 \right] \right\} \\
&= \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=0}^{L} x_i^2 - 2x_i s_i + s_i^2 - x_i^2 \right\} \\
&= \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=0}^{L} s_i^2 - 2x_i s_i \right\} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \\
& \frac{1}{2\sigma^2} \sum_{i=0}^{L} 2x_i s_i - s_i^2 \underset{H_0}{\overset{H_1}{\gtrless}} \ln\gamma \\
& \sum_{i=0}^{L} x_i s_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{1}{2} \left(2\sigma^2 \ln\gamma + \sum_{i=0}^{L} s_i^2 \right).
\end{aligned}$$

2. Sufficient Statistic

In deciding which of two (or more) hypothesis is true based on a set of observations $X_1, ..., X_n$ in many cases the likelihood ratio depends not on each individual X_i but on some function of the observations. For example, in the last example of the previous subsection the likelihood ratio depended only on

$$S(x) = \sum_{i=0}^{L} x_i s_i$$

So if we do not know the individual values of $x_1, ..., x_L$ but know $S(x)$ then that is sufficient information to make the optimal decision about which hypothesis is true.

Definition: Let X be a random variable (vector) whose distribution depends on a parameter $\theta \in \Theta$ (e.g. $\Theta = \{0,1\}$). A function $S(x)$ is said to be sufficient for θ if the conditional distribution of X given $S = s$ is independent of θ.

$$\begin{aligned}
p_\theta(x|s) &= p(x|s), \quad \theta \in \Theta \\
\Leftrightarrow p_\theta(x) &= g_\theta(s(x))h(x) \quad \theta \in \Theta.
\end{aligned}$$

(See Ferguson, pg. 115).

Example:

Consider the repetition code example 2 above where

$$\begin{aligned}
P\{X_i = 1|\theta = 0\} &= \rho \\
P\{X_i = 0|\theta = 0\} &= 1 - \rho \\
P\{X_i = 1|\theta = 1\} &= 1 - \rho \\
P\{X_i = 0|\theta = 1\} &= \rho
\end{aligned}$$

and $\{X_i; i = 1, ..., L-1\}$ are conditionally independent (given θ). The sufficient statistic in this case is $S = \sum_{i=0}^{L-1} X_i$. To see this consider $P\{X_0 = x_0, ..., X_{L-1} = x_{L-1} | S = l, \theta = k\}$ for $k = 0, 1$

$$P\{X_0 = x_0, ..., X_{L-1} = x_{L-1} | S = l, \theta = k\} = \frac{P\{X_0 = x_0, ..., X_{L-1} = x_{L-1}, S = l | \theta = k\}}{P\{S = l | \theta = k\}}$$

$$P\{X_0 = x_0, ..., X_{L-1} = x_{L-1}, S = l | \theta = k\} = \begin{cases} 0, & l \neq \sum_{i=0}^{L-1} x_i \\ p^l (1-p)^{L-l}, & k = 0, l = \sum_{i=0}^{L-1} x_i \\ p^{L-l}(1-p)^l, & k = 1, l = \sum_{i=0}^{L-1} x_i \end{cases}$$

$$P\{S = l | \theta = k\} = \begin{cases} \binom{L}{l} p^l (1-p)^{L-l}, & k = 0 \\ \binom{L}{l} p^{L-l}(1-p)^l, & k = 1. \end{cases}$$

Thus

$$P\{X_0 = x_0, ..., X_{L-1} = x_{L-1} | S = l, \theta = k\} = \begin{cases} 0, & l \neq \sum_{i=0}^{L-1} x_i \\ \binom{L}{l}^{-1}, & l = \sum_{i=0}^{L-1} x_i \end{cases}$$

Thus $P\{X_0 = x_0, ..., X_{L-1} = x_{L-1} | S = l, \theta = k\}$ does not depend on k so S is a sufficient statistic.

3. M-ary Detection Problem

Now consider the problem of deciding which of M hypothesis is true based on observing a random variable. For this case we restrict the performance criteria to be the average error probability. First we write the symbol error probability $P_{e,s}$ in terms of the conditional density functions of the observations as follows.

$$\begin{aligned} P_{e,s} = E[C] &= \sum_{i=0}^{M-1} \left[\sum_{j \neq i} P\{\text{decide } H_j | H_i \text{ true}\} \right] \pi_i \\ &= \sum_{i=0}^{M-1} [1 - P\{\text{decide } H_i | H_i \text{ true}\}] \pi_i \\ &= \sum_{i=0}^{M-1} \pi_i - \sum_{i=0}^{M-1} \int_{R_i} p_i(x) \pi_i dx \\ &= 1 - \sum_{i=0}^{M-1} \int_{R_i} p_i(x) \pi_i dx \ . \end{aligned}$$

We conclude that the decision rule that minimizes average cost assigns x to R_i if $p_i(x)\pi_i = \max_{0 \leq j \leq M-1} p_j(x)\pi_j$.

Thus for M hypotheses the decision rule that minimizes average error probability is to choose i so that $p_i(x)\pi_i > p_j(x)\pi_j, \ \forall \ j \neq i$. Let

$$\Lambda_{i,j} = \frac{p_i(x)}{p_j(x)}$$

where $i = 0, 1, \ldots, M-1$, $j = 0, 1, \ldots, M-1$. Then the optimal decision rule is:

$$\text{Choose } i \text{ if } \Lambda_{i,j} > \frac{\pi_j}{\pi_i} \text{ for all } j \neq i.$$

We will usually assume $\pi_i = \frac{1}{M} \ \forall i$. (If not we should do source encoding to reduce the entropy (rate)). For this case the optimal decision rule is

$$\text{Choose } i \text{ if} \qquad \Lambda_{i,j} > 1 \ \forall \ j \neq i.$$

2. Likelihood Ratio for Random Processes

In this section we explore the extension of the detection process to waveforms (as opposed to finite dimensional random vectors). Consider the simple binary detection problem

$$\begin{aligned} H_0: &\quad x(t) = x_0(t) \\ H_1: &\quad x(t) = x_1(t) \end{aligned}$$

where $x_0(t), x_1(t)$ are random processes with presumably different statistics. For example, $x_0(t)$ could be a (deterministic) signal $s_0(t)$ plus additive white Gaussian noise while $x_1(t)$ could be a signal $s_1(t)$ plus additive white Gaussian noise. The goal is to determine a decision rule based on observing $x(t)$ to decide if H_0 or H_1 is true. The idea is to represent the observation (a random process) using the Karhunen-Loève transform as an infinite sequence of independent random variables.

$$\begin{aligned} H_0: &\quad x(t) = \textstyle\sum_{l=0}^{\infty} x_l \phi_l(t) \quad x_l \text{ has density } p_0(x_l) \\ H_1: &\quad x(t) = \textstyle\sum_{l=0}^{\infty} x_l \phi_l(t) \quad x_l \text{ has density } p_1(x_l). \end{aligned}$$

Here the random variables x_l have different density functions depending on which hypothesis is true. Since the Karhunen-Loève transform represents the random process by a infinite sequence of independent random variables the likelihoods for a finite set of these random variables is a product of the individual conditional density functions for these variables. So consider the likelihood of a finite number of these random variables:

$$p_0^{(N)}(x) = \prod_{l=0}^{N} p_0(x_l).$$

If we let N become large then the likelihood goes to zero. However if we first normalize the product by another density function, say $p^{(N)}(x)$ which also goes to zero and is the same for each hypothesis we can obtain something meaningful when we take the limit of $N \to \infty$. This is illustrated in the following example.

1. Example

Three signals in additive white Gaussian noise For additive white Gaussian noise $K(s,t) = \frac{N_0}{2}\delta(t-s)$. Let $\{\varphi_i(t)\}_{i=0}^{\infty}$ be any complete orthonormal set on $[0,T]$. Consider the case of 3 signals. Find the decision rule to minimize average error probability. First expand the noise using orthonormal set of functions and random variables.

$$n(t) = \sum_{i=0}^{\infty} n_i \varphi_i(t)$$

where $E[n_i] = 0$ and $Var[n_i] = N_0/2$ and $\{n_i\}_{i=0}^{\infty}$ is an independent identically distributed (i.i.d.) sequence of random variables with Gaussian density functions.

Let

$$\begin{aligned} s_0(t) &= \varphi_0(t) + 2\varphi_1(t) \\ s_1(t) &= 2\varphi_0(t) + \varphi_1(t) \\ s_2(t) &= \varphi_0(t) - 2\varphi_1(t) \end{aligned}$$

Note that the energy of each of the three signals is the same, i.e. $\int_0^T s_i^2(t)dt = ||s_i||^2 = 5$. Then we have a three hypothesis testing problem.

$$\begin{aligned} H_0: r(t) &= s_0(t) + n(t) = \sum_{i=0}^{\infty} (s_{0,i} + n_i)\varphi_i(t) \\ H_1: r(t) &= s_1(t) + n(t) = \sum_{i=0}^{\infty} (s_{1,i} + n_i)\varphi_i(t) \\ H_2: r(t) &= s_2(t) + n(t) = \sum_{i=0}^{\infty} (s_{2,i} + n_i)\varphi_i(t) \end{aligned}$$

The decision rule to minimize the average error probability is given as follows

$$\text{Decide } H_i \text{ if } \pi_i p_i(\mathbf{r}) = \max_j \pi_j p_j(\mathbf{r})$$

First let us normalize each side by the density function for the noise alone. The noise density function for $N+1$ variables is

$$p^{(N)}(\mathbf{r}) = \left(\frac{1}{\sqrt{2\pi N_0/2}}\right)^N \exp\{-\frac{1}{2\frac{N_0}{2}}\sum_{i=0}^{N} r_i^2\}$$

The the optimal decision rule is equivalent to

$$\text{Decide } H_i \text{ if } \pi_i \frac{p_i(\mathbf{r})}{p(\mathbf{r})} = \max_j \pi_j \frac{p_j(\mathbf{r})}{p(\mathbf{r})}.$$

As usual assume $\pi_i = 1/M$. Then

$$\begin{aligned}
\frac{p_0^{(N)}(\mathbf{r})}{p^{(N)}(\mathbf{r})} &= \frac{\left(\frac{1}{\sqrt{2\pi N_0/2}}\right)^N \exp\{-\frac{1}{2\frac{N_0}{2}}[\sum_{i=0,1}(r_i - s_{0,i})^2 + \sum_{i=3}^{N} r_i^2]\}}{\left(\frac{1}{\sqrt{2\pi N_0/2}}\right)^N \exp\{-\frac{1}{2\frac{N_0}{2}}\sum_{i=0}^{N} r_i^2\}} \\
&= \exp\{-\frac{1}{N_0}[\sum_{i=0,1}(r_i - s_{0,i})^2 - r_i^2]\} \\
&= \exp\{+\frac{1}{N_0}[2r_1 + 4r_2 - 5]\}.
\end{aligned}$$

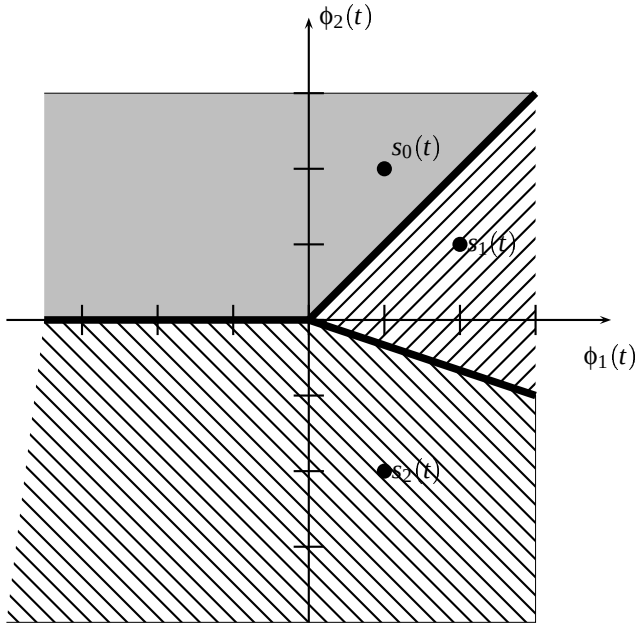
Now since the above doesn't depend on N we can let $N \to \infty$ and the result is the same, i.e.

$$\frac{p_0(\mathbf{r})}{p(\mathbf{r})} \triangleq \lim_{N\to\infty} \frac{p_0^{(N)}(\mathbf{r})}{p^{(N)}(\mathbf{r})} = \exp\{+\frac{1}{N_0}[2r_1 + 4r_2 - 5]\}.$$

Similarly

$$\frac{p_1(\mathbf{r})}{p(\mathbf{r})} = \exp\{+\frac{1}{N_0}[4r_1 + 2r_2 - 5]\}$$

$$\frac{p_2(\mathbf{r})}{p(\mathbf{r})} = \exp\{+\frac{1}{N_0}[2r_1 - 4r_2 - 5]\}.$$



The decision regions are illustrated in the above figure. Note that the decision rule does not depend on the values of $r_3, r_4, \ldots$. These variables are irrelevant when it comes to making an optimal decision. The optimal decision rule is to compute r_1 and r_2 and then find the maximum of $r_1 + 2r_2$, $2r_1 + r_2$, $r_1 - 2r_2$ and then decide H_0 H_1 or H_2 depending on which one is largest. The overall receiver is then shown in the below figure. There are clearly many implicit assumptions in this receiver diagram. One of these assumptions is that we are able to syncronize the correlating waveforms $\phi_l(t)$ to the incoming signal. Another of these is perfect integration over the time span of the waveform. An alternate implementation whereby the multiplier and integrator are replaced with a matched filter is also possible.

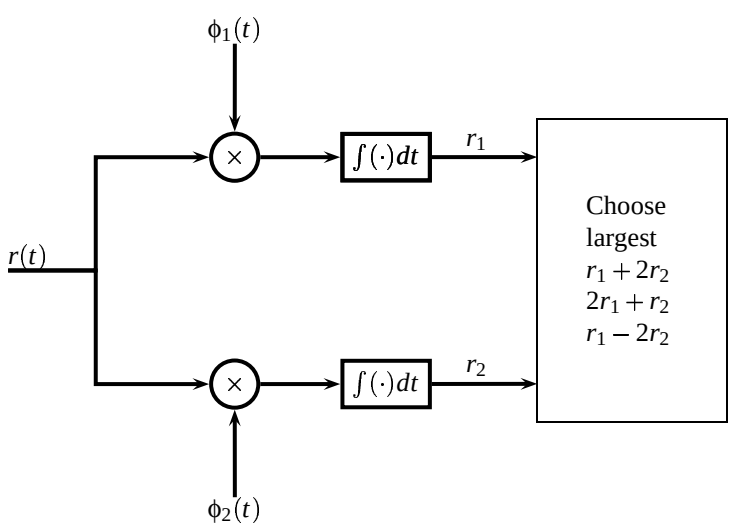


Figure 2.1: Optimum Receiver

3. Detection with Unwanted Parameters

In this section we consider the problem of detection with unwanted parameters. To illustrate consider the problem of minimizing the bit error probability in an M-ary orthogonal signal set.

Let $s_0(t), ..., s_{M-1}(t)$ be orthogonal signals.

$$\begin{aligned} s_0(t) &= \sqrt{E}\phi_0(t) \\ s_1(t) &= \sqrt{E}\phi_1(t) \\ . & \quad . \\ . & \quad . \\ . & \quad . \\ s_{M-1}(t) &= \sqrt{E}\phi_{M-1}(t) \end{aligned}$$

Let $b_0, ..., b_{k-1}$ be the sequence of bits determining which of the M signals is transmitted. Assume the bits are independent and equally likely.

The receiver consists of a bank of matched filters (correlators) that generate a sufficient statistic. If signal s_j is transmitted then

$$\begin{aligned} r_0 &= \delta(j,0)\sqrt{E} + \eta_0 \\ r_1 &= \delta(j,1)\sqrt{E} + \eta_1 \\ . & \quad . \\ . & \quad . \\ . & \quad . \\ . & \quad . \\ r_{M-1} &= \delta(j,M-1)\sqrt{E} + \eta_{M-1} \end{aligned}$$

Consider the detection of data bit b_0. That is, we are interested in minimizing the probability of error for data bit b_0. Let H_0 be the event that $b_0 = 0$ and H_1 be the event that $b_0 = 1$. Let $r = (r_0, r_1, ...r_{M-1})$. Then the optimal receiver must compare the two aposteriori probabilities

$$p(r|H_0)\pi_0 \underset{H_1}{\overset{H_0}{\gtrless}} p(r|H_1)\pi_1$$

To calculate $p(r|H_0)$ we proceed as follows.

$$\begin{aligned} p(r|H_0)\pi_0 &= p(r|b_0 = 0)\pi_0 \\ &= \pi_0 \sum_{b_1,...,b_{k-1}} p(r|b_0 = 0, b_1, ...b_{k-1})p(b_1)p(b_2)\cdots p(b_k) \\ &= 2^{-k} \sum_{b_1,...,b_{k-1}} (\frac{1}{\sqrt{2\pi}\sigma})^M \exp\{-\frac{1}{2\sigma^2}\sum_{l=0}^{M-1}(r_l - s(b))^2\} \\ &= 2^{-k} \sum_{m=0}^{M/2} (\frac{1}{\sqrt{2\pi}\sigma})^M \exp\{-\frac{1}{2\sigma^2}\sum_{l=0}^{M-1}(r_l - \delta(l,m)\sqrt{E})^2\} \\ &= 2^{-k} (\frac{1}{\sqrt{2\pi}\sigma})^M \sum_{m=0}^{M/2} \exp\{-\frac{1}{2\sigma^2}\sum_{l=0}^{M-1}(r_l - \delta(l,m)\sqrt{E})^2\} \\ &= 2^{-k} (\frac{1}{\sqrt{2\pi}\sigma})^M \sum_{m=0}^{M/2} \exp\{-\frac{1}{2\sigma^2}\sum_{l=0}^{M-1}(r_l^2 - 2r_l\delta(l,m)\sqrt{E} + \delta(l,m)E\} \end{aligned}$$

$$
\begin{aligned}
&= 2^{-k}(\frac{1}{\sqrt{2\pi}\sigma})^M \exp\{-\frac{1}{2\sigma^2}\sum_{l=0}^{M-1} r_l^2\}\exp\{-E/2\sigma^2\}\sum_{m=0}^{M/2}\exp\{\frac{1}{\sigma^2}\sum_{l=0}^{M-1}(r_l\delta(l,m)\sqrt{E}\} \\
&= 2^{-k}(\frac{1}{\sqrt{2\pi}\sigma})^M \exp\{-\frac{1}{2\sigma^2}\sum_{l=0}^{M-1} r_l^2\}\exp\{-E/2\sigma^2\}\sum_{m=0}^{M/2}\exp\{\frac{r_m\sqrt{E}}{\sigma^2}\}
\end{aligned}
$$

Similarly

$$
\begin{aligned}
p(r|H_1)\pi_1 &= p(r|b_0=1)\pi_1 \\
&= 2^{-k}(\frac{1}{\sqrt{2\pi}\sigma})^M \exp\{-\frac{1}{2\sigma^2}\sum_{l=0}^{M-1} r_l^2\}\exp\{-E/2\sigma^2\}\sum_{m=M/2+1}^{M-1}\exp\{\frac{r_m\sqrt{E}}{\sigma^2}\}
\end{aligned}
$$

Notice that many of the factors in $p(r|H_1)\pi_1$ and $p(r|H_1)\pi_1$ are the same. Thus the likelihood ratio for bit b_0 is

$$
\frac{p(r|H_1)\pi_1}{p(r|H_0)\pi_0} = \frac{\sum_{m=M/2+1}^{M-1}\exp\{\frac{r_m\sqrt{E}}{\sigma^2}\}}{\sum_{m=0}^{M/2}\exp\{\frac{r_m\sqrt{E}}{\sigma^2}\}}
$$

The log-likelihood ratio is

$$
\log(\frac{p(r|H_1)\pi_1}{p(r|H_0)\pi_0}) = \log(\sum_{m=M/2+1}^{M-1}\exp\{\frac{r_m\sqrt{E}}{\sigma^2}\}) - \log(\sum_{m=0}^{M/2}\exp\{\frac{r_m\sqrt{E}}{\sigma^2}\})
$$

This can be approximated by

$$
\log(\frac{p(r|H_1)\pi_1}{p(r|H_0)\pi_0}) \approx \max_{m=M/2+1}^{M-1}(r_m\sqrt{E}/\sigma^2) - \max_{m=0}^{M/2}(r_m\sqrt{E}/\sigma^2)
$$

4. Likelihood Ratio for Real Signals

In this section we consider the general binary problem of detecting one of two real signals in additive Gaussian noise.

$$
\begin{aligned}
H_0 &: \quad r(t) = s_0(t) + n(t) \\
H_1 &: \quad r(t) = s_1(t) + n(t)
\end{aligned}
$$

Let $n(t)$ have covariance $K(s,t)$ with eigenfunction $\varphi_i(t)$ and eigenvalues $\lambda_i(n(t)$ is also zero mean Gaussian). By $K-L$ expansion $n(t) = \sum_{i=1}^{\infty} n_i\varphi_i(t)$ where n_i is a Gaussian random variable with mean 0 variance λ_i and $E[n_i n_j] = 0 \Rightarrow n_i, n_j$ independent ($n(t)$ is real). Since $\varphi_i(t)$ are a complete orthonormal set and we assume $s_j(t)$ has finite energy we have $s_j(t) = \sum_{i=0}^{\infty} s_{j,i}\varphi_i(t)$. Thus

$$
\begin{aligned}
H_j : r(t) &= \sum_{i=0}^{\infty}(s_{j,i} + n_i)\varphi_i(t) \\
& \quad r_i = s_{j,i} + n_i, \quad i = 1,2,...
\end{aligned}
$$

Define

$$
\Lambda_{j,i}(N) \triangleq \frac{p_j(r_1, r_2, \ldots, r_N)}{p_i(r_1, r_2, \ldots, r_N)} \quad .
$$

$$
\Lambda_{j,i}(r(t)) \triangleq \lim_{N\to\infty}\Lambda_{ji}(N)
$$

where r_i is Gaussian mean $s_{j,i}$ variance λ_i.

$$\begin{aligned}
p_j(r_i) &= \frac{1}{\sqrt{2\pi}\lambda_i}\exp\left\{-\frac{1}{2\lambda_i}(r_i - s_{j,i})^2\right\} \\
p_j(\underline{r}) = \prod_{i=0}^{N} p_j(r_i) &= \prod_{i=0}^{n}(\sqrt{2\pi}\lambda_i)^{-1}\exp\left\{-\frac{1}{2}\sum_{i=0}^{N}\frac{(r_i - s_{j,i})^2}{\lambda_i}\right\} \\
\Lambda_{j,l}(N) = \frac{p_j^N(\underline{r})}{p_l^N(\underline{r})} &= \frac{\prod_{i=0}^{N}(\sqrt{2\pi\lambda_i})^{-1}\ \exp\left\{-\frac{1}{2}\sum_{i=0}^{N}\frac{(r_i - s_{j,i})^2}{\lambda_i}\right\}}{\prod_{i=0}^{N}(\sqrt{2\pi}\lambda_i)^{-1}\ \exp\left\{-\frac{1}{2}\sum_{i=0}^{N}\frac{(r_i - s_{l,i})^2}{\lambda_i}\right\}} \\
&= \exp\left\{-\frac{1}{2}\sum_{i=0}^{N}\frac{1}{\lambda_i}[r_i^2 - 2r_i s_{j,i} + s_{j,i}^2 - r_i^2 + 2r_i s_{l,i} - s_{l,i}^2]\right\} \\
&= \exp\left\{-\frac{1}{2}\sum_{i=0}^{N}\frac{1}{\lambda_i}[s_{j,i}^2 - s_{l,i}^2 + 2r_i(s_{l,i} - s_{j,i})]\right\}.
\end{aligned}$$

Let

$$q_j(t) = \lim_{N\to\infty}\sum_{i=0}^{N}\frac{s_{j,i}}{\lambda_i}\varphi_i(t) = \sum_{i=0}^{\infty}\frac{s_{j,i}\varphi_i(t)}{\lambda_i}$$

Then

$$\begin{aligned}
\int r(t)q_j(t)dt &= \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\int r_l\varphi_l(t)\ \frac{s_{j,i}}{\lambda_i}\varphi_i(t)dt \\
&= \sum_{i=0}^{\infty}\frac{r_i s_{j,i}}{\lambda_i}
\end{aligned}$$

$$\begin{aligned}
\int s_j(t)q_j(t)dt &= \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\int s_{j,i}\varphi_i(t)\frac{s_{j,l}}{\lambda_l}\varphi_l(t)dt \\
&= \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\frac{s_{j,i}s_{j,l}}{\lambda_l}\int\varphi_i(t)\varphi_l(t)dt \\
&= \sum_{i=0}^{\infty}\frac{s_{j,i}^2}{\lambda_i} = (s_j, q_j).
\end{aligned}$$

Thus

$$\Lambda_{j,l}(r(t)) = \lim_{N\to\infty}\Lambda_{j,l}(N) = \exp\left\{-\frac{1}{2}[(s_j,q_j) - (s_l,q_l) + 2(r,q_l) - 2(r,q_j)]\right\}.$$

Note: $q_j(t)$ is solution of the integral equation

$$\int K(s,t)q_j(t)dt = s_j(s)$$

$$\Downarrow$$

$$q_j(t) = \sum_{i=0}^{\infty}\frac{s_{j,i}}{\lambda_i}\ \varphi_i(t).$$

So

$$q_j(s) = \int K^{-1}(s,t)s_j(t)dt$$

$$q_j = K^{-1}s_j$$

If the noise is white, then the noise power in each direction is constant (say λ) and thus

$$q_j(t) = \sum_{i=0}^{\infty} \frac{s_{j,i}}{\lambda} \varphi_i(t) = \frac{1}{\lambda} s_j(t).$$

The optimal receiver then becomes

$$\Lambda_{j,l}(r(t)) = \exp\left\{ -\frac{1}{2\lambda}[(s_j, s_j) - (s_l, s_l) + 2(r, s_l) - 2(r, s_j)] \right\}.$$

or equivalently

$$\Lambda_{j,l}(r(t)) = \exp\left\{ -\frac{1}{2\lambda}[||s_j||^2 - ||s_l||^2 + 2(r, s_l - s_j)] \right\}.$$

For equal energy signals this amounts to picking the signal with the largest correlation with the received signal.

The optimal receiver in nonwhite Gaussian noise can be implemented in a similar fashion as shown below.

$$\begin{aligned} (s_j, q_j) &= (s_j, K^{-1} s_j) = (K^{-1/2} s_j, K^{-1/2} s_j) = \|K^{1/2} s_j\|^2 \\ (r, q_j) &= (r, K^{-1} s_j) = (K^{-1/2} r, K^{-1/2} s_j) \end{aligned}$$

Thus

$$\Lambda_{j,l}(r(t)) = \exp\left\{ -\frac{1}{2}[||K^{-1/2} s_j||^2 - ||K^{-1/2} s_l||^2 + 2(K^{-1/2} r, K^{-1/2}(s_l - s_j))] \right\}.$$

It is clear then that this is just the optimal filter for signals $K^{-1/2} s_j$ when received in additive white Gaussian noise. This approach is called "whitening" because $K^{-1/2} n$ will be a white Gaussian noise process.

5. Likelihood Ratio for Complex Signals

In this section we rederive the likelihood ratios for complex signals received in complex noise. We assume that the signals are the lowpass representation of bandpass signal and the noise is the lowpass representation of a narrowband random process. Let

$$\begin{aligned} H_0 : r(t) &= s_0(t) + n(t) \\ H_1 : r(t) &= s_1(t) + n(t) \end{aligned}$$

where $n(t)$ has covariance $K(s,t)$, with eigenfunctions $\varphi_i(t)$, eigenvalues λ_i. Using $K-L$ expansion we have

$$H_i : r(t) = \sum_{j=0}^{\infty} (s_{i,j} + n_j)\varphi_j(t)$$

$$r_j = s_{i,j} + n_j.$$

$$\begin{aligned} \frac{p_j(r_1,\ldots,r_n)}{p_i(r_1,\ldots,r_n)} &= \frac{\prod_{l=0}^{n} \frac{1}{\pi\lambda_l}\, e^{-(|r_l - s_{j,l}|^2)/\lambda_l}}{\prod_{l=0}^{n} \frac{1}{\pi\lambda_l}\, e^{-(|r_l - s_{i,l}|^2)/\lambda_l}} \\ &= \prod_{l=0}^{n} \exp\{-(|r_l - s_{j,l}|^2 - |r_l - s_{il}|^2)/\lambda_l\} \\ &= \exp\left\{-\sum_{l=0}^{n} (|r_l - s_{j,l}|^2 - |r_l - s_{i,l}|^2)/\lambda_l\right\} \\ &= \exp\left\{-\sum_{l=0}^{n} \frac{|r_l|^2 + |s_{j,l}|^2 - 2\,\mathrm{Re}\,(r_l s_{j,l}^*) - |r_l|^2 - |s_{i,l}|^2 + 2\,\mathrm{Re}\,(r_l s_{i,l}^*)}{\lambda_l}\right\} \\ &= \exp\left\{-\sum_{l=0}^{n} \frac{|s_{j,l}|^2 - |s_{i,l}|^2 + 2\mathrm{Re}\,[r_l(s_{i,l} - s_{j,l})^*]}{\lambda_l}\right\}. \end{aligned}$$

Let $q_j(t) = \sum_{i=0}^{\infty} \frac{s_{ji}}{\lambda_i}\varphi_i(t)$ then

$$\begin{aligned} (q_j(t), s_j(t)) &= \int_a^b \sum_{l=0}^{\infty} \frac{s_{j,l}}{\lambda_l}\varphi_l(t) \sum_{k=0}^{\infty} s_{j,k}^* \varphi_k^*(t)dt \\ &= \sum_{l=0}^{\infty} |s_{j,l}|^2/\lambda_l = (s_j(t), q_j(t)) \end{aligned}$$

$$\begin{aligned} (r(t), q_j(t)) &= \int_a^b \sum_{l=0}^{\infty} r_l\varphi_l(t) \sum_{k=0}^{\infty} \frac{s_{j,k}^*}{\lambda_k}\, \varphi_k^*(t)dt \\ &= \sum_{l=0}^{\infty} \frac{r_l s_{j,l}^*}{\lambda_l}. \end{aligned}$$

So

$$\begin{aligned} \Lambda_{ji}(r(t)) &= \lim_{H\to\infty} \Lambda_{ji}^{(n)}(r(t)) = \frac{p_j(r_1,\ldots,r_n)}{p_i(r_1,\ldots,r_n)} \\ &= \exp\{-[(s_j,q_j) - (s_i,q_i) + 2\,\mathrm{Re}\,(r(t), q_i(t) - q_j(t))] \end{aligned}$$

Note: Since we are dealing with noise that is derived from a narrowband random process we can not use the results derived for *real* random processes we *must* use the likelihood ratio for *complex* random process given above.

For real random process the likelihood ratio is

$$\Lambda_{ji}(N) = \exp\{-\frac{1}{2}[(s_j,q_j)-(s_i,q_i)+2(r,q_i-q_j)]\}.$$

For additive white Gaussian noise (real)

$$\begin{aligned} q_i(t) = \sum_{j=1}^{\infty} \frac{s_{i,j}\varphi_j(t)}{\lambda_j} &= \frac{2}{N_0}\sum_{j=1}^{\infty} s_{i,j}\varphi_j(t) \\ &= \frac{2}{N_0}s_i(t) \end{aligned}$$

So the likelihood ratio (for real signals) becomes

$$\begin{aligned} \Lambda_{j,l} &= \lim_{N\to\infty}\frac{p_j(r_1,\ldots,r_N)}{p_l(r_1,\ldots,r_N)} = \exp\left\{-\frac{1}{2}\left[\frac{2}{N_0}((s_j,s_j)-(s_l,s_l))+2\cdot\frac{2}{N_0}(r,s_l-s_j)\right]\right\} \\ &= \exp\left\{-\frac{1}{N_0}[\|s_j\|^2-2(r,s_j)-\|r\|^2-(\|s_l\|^2-2(r,s_l)-\|r\|^2)]\right\} \\ &= \exp\left\{-\frac{1}{N_0}[\|s_j-r\|^2-\|s_l-r\|^2]\right\} \underset{H_l}{\overset{H_j}{\gtrless}} 1. \end{aligned}$$

Assume $\pi_j = \frac{1}{M} \quad j = 1,2,\ldots,M$. Then $\alpha = 1$. An equivalent decision rule then is

$$\|s_j-r\|^2-\|s_l-r\|^2 \underset{H_j}{\overset{H_l}{\gtrless}} 0$$

$$\|s_j-r\|^2 \underset{H_j}{\overset{H_l}{\gtrless}} \|s_l-r\|^2.$$

The optimum decision rule for additive white Gaussian noise is then to choose i if

$$\|s_i-r\|^2 = \min_{1\le j\le M}\|s_j-r\|^2.$$

6. Example: M orthogonal signals in additive white Gaussian noise

In this section we consider the optimum receiver for M-ary orthogonal signals and the associated error probaiblity. Assume the M signals are equienergy signals and equiprobable. The decision rule derived previously for AWGN is

$$\text{Decide } H_i \text{ if } ||s_i-r||^2 = \min_{1\le j\le M}||s_j-r||^2.$$

Now since the M signals are orthogonal and equienergy we can write this as

$$||s_j-r||^2 = ||s_j||^2-2(s_j,r)+||r||^2.$$

The first term above is constant for each j as is the last term. Thus finding the minimum is equivalent to finding the maximum of

$$(s_j,r).$$

Thus the receiver should compute the inner product between the M different signals and find the largest such correlation. If the signals are all of duration T, i.e. zero outside the interval $[0,T]$ then this is also equivalent to filtering the received signal with a filter with impluse response $s_j(T-t)$, sampling the output of the filter at time T and choosing the largest as shown below.

7. Problems

1. (a) Consider any two binary signals $s_0(t), s_1(t)$ with $E_i = \| s_i(t) \|^2$ the energy of signal $s_i(t)$ and

$$\begin{aligned} \rho &= (s_0, s_1)/\bar{E} \\ &= \int s_0(t) s_1(t) dt / \bar{E} \end{aligned}$$

where $\bar{E} = (E_0 + E_1)/2$. Assume that signal $s_0(t)$ is transmitted with probability π_0 and signal $s_1(t)$ is transmitted with probability π_1 with $\pi_0 + \pi_1 = 1$. The received signal is the transmitted signal plus additive white Gaussian noise with two-sided power spectral density $N_0/2$.

(a) Find the optimum receiver. (Do **not** assume that $\pi_0 = \pi_1$).

(b) Find the average error probability of the optimum receiver.

(c) If now $\pi_0 = \pi_1$ find the value of ρ that minimizes the error probability. (Note that by Schwarz' inequality $-1 \leq \rho \leq 1$).

2. Consider the following hypothesis testing problem

$$H_0 : X \quad p_0(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-(x-1)^2/2\right\}$$

$$H_1 : X \quad p_1(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-(x+1)^2/2\right\}$$

(a) Find the test that minimizes the average cost.

(b) If $c_{00} = c_{11} = 0$ and $c_{01} = c_{10} = 1$ and $\pi_0 = \pi_1$ what is the (minimum average cost) test?

3. Consider a hypothesis testing problem with three alternatives $H_0, H_1,$ and H_2. With c_{ij} representing the cost of deciding H_i is true when H_j is true, find a test(s) that minimizes the average cost. (Let $\pi_i \;\; i = 0, 1, 2$ be the probability that H_i is true).

4. Find the test (decision rule) that minimizes average error probability for the following detection theory problem. Assume $\sigma_1 > \sigma_0$ and $\pi_0 = \pi_1 = 1/2$.

$$H_0 : X \quad p_0(x) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left\{-x^2/(2\sigma_0^2)\right\}$$

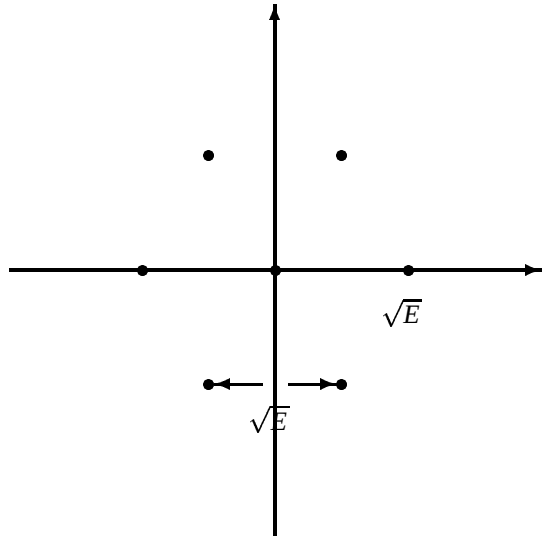
$$H_1 : X \quad p_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-x^2/(2\sigma_1^2)\right\}$$

5. For the seven-signal set shown, transmitted over the additive white Gaussian noise (AWGN) channel, with equal a priori probabilities ($\pi_i = 1/7$) and equal energies (for the six nonzero signals).

(a) Determine the optimum decision regions.

(b) Show that $P_m =$ probability of error given signal m transmitted is upper bounded by the probability that the norm of the two dimensional noise vector is greater than $\sqrt{E}/2$.

(c) Using (b) find an upper bound on the average probability of error.



6. A certain digital communication system transmits one of two signals s_0, and s_1 in additive (not necessarily white) Gaussian noise with covariance function $K(s,t)$. Let ϕ_i be the eigenfunctions of $K(s,t)$ and let λ_i be the corresponding eigenvalues.

   (a) If $s_0(t) = A_0\phi_0(t)$ and $s_1(t) = A_1\phi_1(t)$ find the optimal receiver for minimizing the average error probability.

   (b) If the two signals are equally probable find the average error probability.

7. A certain communication system transmits one of four equally likely signals in additive WHITE Gaussian noise with power spectral density $N_0/2$. The signals are

$$s_0(t) = A\phi_1(t) + A\phi_2(t)$$

$$s_1(t) = A\phi_1(t) - A\phi_2(t)$$

$$s_2(t) = -A\phi_1(t) + A\phi_2(t)$$

$$s_3(t) = -A\phi_1(t) - A\phi_2(t)$$

   where $\phi_1(t)$ and $\phi_2(t)$ are orthonormal signals. (a) If $r(t)$ is the received signal write down the likelihood ratio for deciding between $s_1(t)$ and $s_0(t)$, the likelihood ratio for deciding between $s_2(t)$ and $s_0(t)$, and the likelihood ratio for deciding between $s_3(t)$ and $s_0(t)$. (b) Using the results in (a) find the decision region R_0 for deciding H_0 (that $s_0(t)$ was transmitted).

8. One of two equally likely messages is to be transmitted over an AWGN channel by means of the two signals

$$s_0(t) = \sqrt{2E_s/T}\,p_T(t)\cos\omega_c t$$

$$s_1(t) = \sqrt{2E_s/T}\,p_T(t)\cos(\omega_c t + \phi)$$

   where $p_T(t)$ is a unit amplitude pulse of duration T.

   (a) Find the optimum receiver. (You may assume that $\omega_c T$ is an integer multiple of π.)

   (b) Calculate the error probability in terms of the Q function.

9. Consider an additive Gaussian channel with positive definite covariance function $K(s,t)$. As in class define the operator $K^{-1/2}(s,t)$ by

$$K^{-1/2}(s,t) = \sum_{i=1}^{\infty} \frac{1}{\sqrt{\lambda_i}}\phi_i(s)\phi_i^*(t)$$

(a) Show that for any two functions $x(t)$, $y(t)$ (in $L_2[a,b]$) that

$$(x, K^{-1}y) = (K^{-1/2}x, K^{-1/2}y)$$

(b) Consider a signal set $\{s_j : j = 0,1,\ldots,M-1\}$. Let $\hat{s}_j = K^{-1/2}s_j$, $j = 0,1,\ldots,M-1$ and let $\hat{r} = K^{-1/2}r$. Consider the receiver which decides hypothesis H_i if

$$||\hat{s}_i - \hat{r}|| = \min_{0\le j\le M-1} ||\hat{s}_j - \hat{r}||$$

Using (a) (or otherwise) show that the above receiver is optimal for equiprobable signals.

10. Consider the binary symmetric channel with crossover probability $p < 1/2$. The input and output alphabet for this channel is $\{0,1\}$. Consider a signal set $\{\mathbf{s}_i : 0 \le i \le M-1\}$ in n dimensions for this channel with each coefficient (component) being 0 or 1, i.e. consider the signal as a vector of length n. Assume that all signals are equally likely. Let $\mathbf{r} = (r_1,\ldots,r_n)$ be the received vector with $r_i \in \{0,1\}$. For any two binary vectors $\mathbf{x}$ and $\mathbf{y}$ let $d_H(\mathbf{x},\mathbf{y})$ = number of places where $\mathbf{x}$ and $\mathbf{y}$ differ.

    (a) Show that the optimal receiver for minimizing average error probability is to chose that signal for which differs from $\mathbf{r}$ in the fewest places i.e. decide hypothesis H_i if

    $$d_H(\mathbf{r},\mathbf{s}_i) = \min_{0\le j\le M-1} d_H(\mathbf{r},\mathbf{s}_j)$$

    (b) For any such signal set determine the union-Bhattacharrya bound for the optimal receiver. Your answer should depend on the distance between signals.

11. Consider a digital communication system that transmits one of two equally likely signals over a discrete time additive channel with nonwhite Gaussian noise. The two signals transmitted are

    $$s_0 = (1,1,-1)$$

    and

    $$s_1 = (1,-1,1).$$

    The noise added to each component of the signal is Gaussian but not independent nor identically distributed. The noise vector $n = (n_0,n_1,n_2)$ has zero mean and covariance matrix.

    $$K = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

    The covariance matrix has eigenvalues, $\lambda_1 = 5$, $\lambda_2 = 5$ and $\lambda_3 = 1$ with corresponding (orthonormal) eigenvectors

    $$\begin{aligned} \phi_1 &= (-1/\sqrt{2}, 1/\sqrt{2}, 0), \\ \phi_2 &= (0,0,1), \\ \phi_3 &= (1/\sqrt{2}, 1/\sqrt{2}, 0). \end{aligned}$$

    (a) Find the optimal receiver for minimizing the error probability between the two signals.

    (b) Find the error probability for the optimal receiver.

12. Consider the following detection problem. Under hypothesis H_0, X is zero mean Gaussian with variance σ^2. Under hypothesis H_1 is a mixture of two Gaussians with means +1 and -1, each with variance σ^2 and the mixture is uniform. That is under H_1, $X = \alpha + \eta$ where α is a random variable equally likely to be ± 1 and η is a zero mean Gaussian random variable. Find the optimal decision region for deciding which of the two equally likely hypothesis is true in order to minimize the error probability.