STOCHASTIC SIGNALS AND PROCESSES

Lecture 3: Detection theory
Signal detection

Goal: Extraction of information contained in signals that are mixed with noise.

Example: Determine the sequence of binary digits over a communication channel

Two classes of information extraction algorithms:
Signal detection and signal estimation
Signal detection

A priori knowledge of the set of symbols or waveforms on both sides.
Not known: What and when
Signal estimation

Determining each value over time

Example: voice signal transmitted over the telephone line

No apriori knowledge of symbols
Binary decision

A source emitting 2 possible outputs at various instants of time: No Disease ($H_0$) or Disease ($H_1$).

Null hypothesis
Alternate hypothesis
Statistical decision

Suppose we have a decision based on a **single observation** $Y$ is a random variable of the outcomes.
The range of values of $Y$ form the observation space $Z$.
$Z = Z_1 \cup Z_0$
Statistical decision

\(f_{Y|H_0}(y|H_0)\) and \(f_{Y|H_1}(y|H_1)\): The pdf of \(Y\) corresponding to each hypothesis. For each binary decision (\(D_0\) and \(D_1\)) we have 4 possible cases:

<table>
<thead>
<tr>
<th>(D_0)</th>
<th>(D_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>Type 1 error: False alarm</td>
</tr>
<tr>
<td>Type 2 error: Miss detection</td>
<td>Detection</td>
</tr>
</tbody>
</table>
How to solve decision problems?

Bayes’ Criterion

Minimax Criterion*

Neyman – Pearson Criterion*

* SELF STUDY
Bayes’ criterion

In Bayes’ criterion, 2 assumptions are made:
1. The probability of the two source outputs is known, $P(H_0)$ or $P_0$ and $P(H_1)$ or $P_1$: a priori probabilities. $P_0 + P_1 = 1$

2. A cost function is assigned to each possible decision.

Example: Consequence of starting treatment whether a patient has cancer or No
Bayes’ criterion

Let define $D_i, i = 0, 1$
where $D_0 \Rightarrow$ decide $H_0$ and $D_1 \Rightarrow$ decide $H_1$

We can define $C_{ij}, i and j = 0, 1$. the cost associated
to decision $D_i$, given that the true hypothesis is $H_j$.

<table>
<thead>
<tr>
<th></th>
<th>$H_0$</th>
<th>$H_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_0$</td>
<td>$C_{00}$</td>
<td>$C_{01}$</td>
</tr>
<tr>
<td>$D_1$</td>
<td>$C_{10}$</td>
<td>$C_{11}$</td>
</tr>
</tbody>
</table>
Bayes’ criterion

Goal: Determine the decision rule so that the average cost $E(C)$ (or Risk) is minimized.

$$R = E(C) = C_{00}P(D_0, H_0) + C_{01}P(D_0, H_1) + C_{10}P(D_1, H_0) + C_{11}P(D_1, H_1)$$

It is also assumed that $C_{01} > C_{11}$ and $C_{10} > C_{00}$

$P(D_i, H_j)$ is the joint probability that we decide $D_i$, and the hypothesis $H_j$ is true.
Baye’s rule states that

\[ P(D_i, H_j) = P(D_i | H_j)P(H_j) \]

The conditional probabilities are

\[ P(D_0 | H_0) = P(\text{decide } H_0 | H_0 \text{ true}) = \int_{Z_0} f_{Y|H_0}(y | H_0) dy \]

\[ P(D_0 | H_1) = P(\text{decide } H_0 | H_1 \text{ true}) = \int_{Z_0} f_{Y|H_1}(y | H_1) dy \]

\[ P(D_1 | H_0) = P(\text{decide } H_1 | H_0 \text{ true}) = \int_{Z_1} f_{Y|H_0}(y | H_0) dy \]

\[ P(D_1 | H_1) = P(\text{decide } H_1 | H_1 \text{ true}) = \int_{Z_1} f_{Y|H_1}(y | H_1) dy \]
We call:

- The probability of detection or $P_D$
- The probability of miss or $P_M$
- The probability of false alarm or $P_F$

<table>
<thead>
<tr>
<th></th>
<th>$H_0$</th>
<th>$H_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_0$</td>
<td>$P(D_0, H_0)$</td>
<td>$P(D_0, H_1)$</td>
</tr>
<tr>
<td>$D_1$</td>
<td>$P(D_1, H_0)$</td>
<td>$P(D_1, H_1)$</td>
</tr>
</tbody>
</table>
Looking at each hypothesis, we see that

\[ P_M = 1 - P_D \]

and

\[ P(D_0 | H_0) = 1 - P_F \]

Now the probability of correct detection will be:

\[
P(\text{correct decision}) = P(c) = P(D_0, H_0) + P(D_1, H_1) \\
= P(D_0 | H_0)P(H_0) + P(D_1 | H_1)P(H_1) \\
= (1 - P_F)P_0 + P_D P_1
\]

And the error

\[
P(\text{error}) = P(\varepsilon) = P(D_0, H_1) + P(D_1, H_0) \\
= P(D_0 | H_1)P(H_1) + P(D_1 | H_0)P(H_0) \\
= P_M P_1 + P_F P_0
\]
The average cost becomes

\[
\mathcal{R} = E[C] = C_{00} P(D_0, H_0) + C_{01} P(D_0, H_1) + C_{10} P(D_1, H_0) + C_{11} P(D_1, H_1)
\]

\[
\mathcal{R} = E[C] = C_{00} (1 - P_F) P_0 + C_{01} (1 - P_D) P_1 + C_{10} P_F P_0 + C_{11} P_D P_1
\]

\[
\mathcal{R} = P_0 C_{00} \int_{Z_0} f_{Y|H_0}(y | H_0) \, dy + P_1 C_{10} \int_{Z_1} f_{Y|H_0}(y | H_0) \, dy + P_0 C_{11} \int_{Z_1} f_{Y|H_1}(y | H_1) \, dy
\]

\[
+ \quad P_1 C_{11} \int_{Z_1} f_{Y|H_1}(y | H_1) \, dy
\]
\[ R = P_0 C_{00} \int_{Z_0} f_{y|H_0}(y \mid H_0) dy + P_1 C_{01} \int_{Z_0} f_{y|H_1}(y \mid H_1) dy + P_0 C_{10} \int_{Z_1} f_{y|H_0}(y \mid H_0) dy + P_1 C_{11} \int_{Z_1} f_{y|H_1}(y \mid H_1) dy \]

\[ R = P_0 C_{10} + P_1 C_{11} + \int_{Z_0} \{ [P_1 (C_{01} - C_{11}) f_{y|H_1}(y \mid H_1)] - [P_0 (C_{10} - C_{00}) f_{y|H_0}(y \mid H_0)] \} dy \]

\[ \int_{Z} f_{y|H_0}(y \mid H_0) dy = \int_{Z} f_{y|H_1}(y \mid H_1) dy = 1 \]
\[ R = P_0 C_{10} \int_{z_0} f_{y_1|H_0} (y_{1|H_0}) \, dy + P_1 C_{01} \int_{z_0} f_{y_1|H_1} (y_{1|H_1}) \, dy + P_0 C_{10} \int_{z_1} f_{y_1|H_0} (y_{1|H_0}) \, dy + P_1 C_{11} \int_{z_1} f_{y_1|H_1} (y_{1|H_1}) \, dy \]

We use
\[ \int_{z_1} f_{y_1|H_0} (y_{1|H_0}) \, dy = 1 - \int_{z_0} f_{y_1|H_0} (y_{1|H_0}) \, dy \]

and
\[ \int_{z_1} f_{y_1|H_1} (y_{1|H_1}) \, dy = 1 - \int_{z_0} f_{y_1|H_1} (y_{1|H_1}) \, dy \]

\[ R = P_0 C_{10} \int_{z_0} f_{y_1|H_0} (y_{1|H_0}) \, dy + P_1 C_{01} \int_{z_0} f_{y_1|H_1} (y_{1|H_1}) \, dy \\
+ P_0 C_{10} \left[ 1 - \int_{z_0} f_{y_1|H_0} (y_{1|H_0}) \, dy \right] + P_1 C_{11} \left[ 1 - \int_{z_0} f_{y_1|H_1} (y_{1|H_1}) \, dy \right] \]

\[ R = P_0 C_{10} + P_1 C_{01} + P_0 \left[ C_{00} - C_{10} \right] \int_{z_0} f_{y_1|H_0} (y_{1|H_0}) \, dy + \\
P_1 \left[ C_{01} - C_{11} \right] \int_{z_0} f_{y_1|H_1} (y_{1|H_1}) \, dy \]

\[ = P_0 C_{10} + P_1 C_{01} + \int \left\{ P_1 \left[ C_{01} - C_{11} \right] f_{y_1|H_1} (y_{1|H_1}) - P_0 \left[ C_{10} - C_{00} \right] f_{y_1|H_0} (y_{1|H_0}) \right\} \, dy \]
because we assumed $C_{01} > C_{11}$ and $C_{10} > C_{00}$

So to minimize the risk, this difference should go towards minus infinity ($-\infty$)
We assign to the region $Z_0$, those points for which

$$P_1(C_{01} - C_{11}) f_{Y|H_1}(y|H_1) < P_0(C_{10} - C_{00}) f_{Y|H_0}(y|H_0)$$

So if

$$P_1(C_{01} - C_{11}) f_{Y|H_1}(y|H_1) > P_0(C_{10} - C_{00}) f_{Y|H_0}(y|H_0)$$

then we decide $H_1$

Thus the decision rule resulting from Bayes’ criterion is

$$\frac{f_{Y|H_1}(y|H_1)}{f_{Y|H_0}(y|H_0)} \xleftarrow{H_1} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \xleftarrow{H_0} \frac{f_{Y|H_1}(y|H_1)}{f_{Y|H_0}(y|H_0)}$$
This is called the *likelihood ratio* and is denoted as

\[
\Lambda(y) = \frac{f_{y|H_1}(y|H_1)}{f_{y|H_0}(y|H_0)}
\]

\[
\eta = \frac{P_0 (C_{10} - C_{00})}{P_1 (C_{01} - C_{11})}
\]

Is called the threshold

Baye’s criterion that minimizes the average cost results in a likelihood ratio test

\[
\begin{align*}
\Lambda(y) &> \eta \\
\text{i.e.} & \\
\ln \Lambda(y) &> \ln \eta
\end{align*}
\]

OR
**Special cases**

Costs of errors equal 1. $C_{01} = C_{10} = 1$
And costs of correct decision equal 0. $C_{00} = C_{11} = 0$.

$$\mathcal{R} = E[C] = C_{00}(1 - P_F)P_0 + C_{01}(1 - P_D)P_1 + C_{10}P_F P_0 + C_{11}P_D P_1$$

$$\mathcal{R} = P_M P_1 + P_F P_0 = P(\varepsilon)$$

Thus minimizing the average cost is equivalent to minimizing the probability of error: *Minimum probability of error receivers*

The threshold becomes:

$$\eta = \frac{P_0 (C_{10} - C_{00})}{P_1 (C_{01} - C_{11})}$$

For equal marginal probabilities $P_1 = P_0 = 0.5$, log likelihood ratio is compared to zero.
Example 5.1

In a digital communication system, consider a source whose output under hypothesis $H_1$ is a constant voltage of value $m$, while its output under $H_0$ is zero. The received signal is corrupted by $N$, an additive white Gaussian noise of zero mean, and variance $\sigma^2$.

(a) Set up the likelihood ratio test and determine the decision regions.

(b) Calculate the probability of false alarm and probability of detection

Solution in the book
M-ary Hypothesis testing

We now consider the choice of one hypothesis among $M$ hypotheses, $H_0, H_1, ..., H_{M-1}$. The risk becomes

$$\mathcal{R} = \sum_{i=1}^{M-1} \sum_{j=1}^{M-1} P_j C_{ij} P(D_i \mid H_j)$$  \hspace{1cm} (5.33)

$$\mathcal{R} = E[C] = C_{00} P(D_0, H_0) + C_{01} P(D_0, H_1) + C_{10} P(D_1, H_0) + C_{11} P(D_1, H_1)$$  \hspace{1cm} (5.5)
Remember the Joint probability

\[ P(D_i, H_j) = P(D_i \mid H_j)P(H_j) \]

Recall also that

\[ P(D_i \mid H_j) = \int_{Z_i} f_{Y\mid H_j}(y \mid H_j) \, dy \]

We have M subspaces

\[ Z = Z_0 \cup Z_1 \cup \ldots \cup Z_{M-1} \]

\[ R = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} P_{ij} C_{ij} \int_{Z_i} f_{Y\mid H_j}(y \mid H_j) \, dy \]

\[ \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} P_{ij} C_{ij} \int_{Z_i} f_{Y\mid H_j}(y \mid H_j) \, dy + \sum_{i=1}^{M-1} P_{ii} C_{ii} \int_{Z_i} f_{Y\mid H_i}(y \mid H_i) \, dy \]
We would like to minimize the integral

\[
I_i(y) = \sum_{j=0}^{M-1} P_j (C_{ij} - C_{jj}) f_{y|H_j} (y | H_j)
\]
Defining the likelihood ratio \( \Lambda_i(y) \), \( i = 1, 2, \ldots, M - 1 \), as

\[
\Lambda_i(y) = \frac{f_{Y|H_i}(y|H_i)}{f_{Y|H_0}(y|H_0)} , \quad i = 1, 2, \ldots, M - 1
\]

Let normalize with the null hypothesis

\[
J_i(y) = \frac{I_i(y)}{f_{Y|H_0}(y|H_0)} = \frac{\sum_{j=0}^{M-1} P_j (C_{ij} - C_{jj}) f_{Y|H_j}(y|H_j)}{f_{Y|H_0}(y|H_0)} = \sum_{j=1}^{M-1} P_j (C_{ij} - C_{jj}) \Lambda_j(y)
\]

Choose the hypothesis for which \( J_i(y) \) is minimized
Maximum a posteriori probability (MAP)

In many problems it is common to have

\[ C_{ii} = 0, \quad i = 1, 2, \ldots, M - 1 \quad C_{ij} = 1, \quad i \neq j \text{ and } i, j = 0, 1, \ldots, M - 1 \]

Thus minimizing the risk = Minimizing the probability of error

\[
I_i(y) = \sum_{\substack{j=0 \quad j \neq i \quad j=0 \quad j \neq i \quad M-1 \quad j=0 \quad j \neq i \quad M-1}} P_j \left( C_{ij} - C_{jj} \right) f_{Y|H_j}(y \mid H_j)
\]

\[
I_i(y) = \sum_{j=0}^{M-1} P(H_j) f_{Y|H_j}(y \mid H_j) = \sum_{j=0}^{M-1} P(H_j \mid Y) f_Y(y) = [1 - P(H_i \mid Y)] f_Y(y)
\]
Maximum a posteriori probability (MAP)

\[ P(A|B) = \frac{P(B|A)P(A)}{P(B)} \]

Thus minimizing \( I_i(y) \) = Maximizing \( P(H_i | Y) \)

\( P(H_i | Y) \)

the a posteriori probability of \( H_i \) given the observation vector \( y \).

The receiver decides in favor of the hypothesis with largest \( P(H_i | Y) \)

This is a Maximum a posteriori probability (MAP) receiver
Detection theory applied to waveforms

A particularly important topic in detection theory. In binary detection we have

\[ H_1: Y(t) = s(t) + w(t), \quad 0 \leq t \leq T \]
\[ H_0: Y(t) = w(t), \quad 0 \leq t \leq T \]

\( s(t) \) a deterministic signal of known waveform
\( w(t) \) additive zero mean WGN

The Energy of the signal is given as

\[ E = \int_0^T s^2(dt) \]
Matched filter

An important topic in detection theory
To filter out noise we apply $K$ matched filters with impulse response.

$$h_k(t) = \Phi_k(T - t), \quad 0 \leq t \leq T. \; k = 1, 2 \ldots K$$

$$\{\Phi_k(t)\} \text{ form the set of basis functions.}$$
Matched filter

If we select $h(t) = s(T - t)$ a time-reversed and delayed version of $s(t)$, we talk about a filter that is matched to the signal.

A way to judge the quality of a filter is to measure the signal-to-noise ratio (SNR)

$$\max \eta, \text{ where } \eta \text{ is peak pulse SNR}$$

$$\eta = \frac{|g_0(T)|^2}{E\{n^2(t)\}} = \frac{\text{instantaneous power}}{\text{average power}}$$
Example 10.5

Let \( s_1(t) \) and \( s_2(t) \) be two signals as shown in figure below, which are used to transmit a binary sequence.

a. Sketch the matched filters
b. Determine and sketch the response to \( s_2(t) \) of the matched filter