EE 631: Estimation and Detection
Part 1

Dr. Mehrdad Soumekh

1. DETECTION

- Testing of hypotheses.
- Statistical decision theory.
- Classification.

Hypothesis

A hypothesis is a statement or claim about the state of nature. In this case the user gathers data and lets the data support or cast doubt on a hypothesis. This is called "testing of hypothesis."
This process involves choosing among a set of actions - a decision problem.

Applications

- Data communications: e.g. "0" or "1" is transmitted from a station to a user. The received signal, which may be contaminated by the channel noise, is used to determine whether "0" or "1" is transmitted.
- Radar/sonar: Decide presence or absence of targets.
- Pattern recognition: Observe a signal and classify it.

Approach

Probabilistic approach required.
Requirements:

- Source model.
- Observation model.
- Criterion for making a decision.

Example: Binary hypothesis testing

R: measurement
N: noise (random)
S: source parameters (random/non-random)

The decision problem is to select between the following two hypotheses:

$$H_0 : R = N \tag{1}$$

(Null hypothesis) when 0 is sent.

$$H_1 : R = N + S \tag{2}$$

when 1 is sent.

This problem could be generated in the following analog system:
A user receives an amplitude shift keying (ASK) signal of the form:

$$r(t) = s(t) + n(t) \tag{3}$$

where

$$s(t) = \begin{cases} 0 & ;\text{ under } H_0 \\ p(t) & ;\text{ under } H_1 \end{cases}$$

$p(t)$ is deterministic.

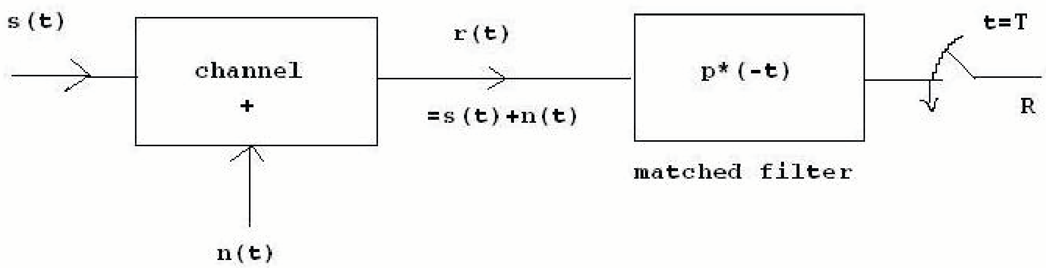

Figure 1. Block diagram of the ASK communication system

$$R = r(t) * p^*(-t)\big|_{t=T} \tag{4}$$

$$N = n(t) * p^*(-t)\big|_{t=T} \tag{5}$$

$$S = s(t) * p^*(-t)\big|_{t=T} \tag{6}$$

where $p^*(-t)$ denotes the matched filter matched to $p(t)$.

2. ESTIMATION

The problem of estimation arises from the need to know the actual value of a signal or its parameters by observing related signals.

Approach

Probabilistic approach required.

Requirements:

- Source model.
- Observation model.
- Criterion (bounds on the moments of the estimator error).

Example

$$r(t) = A\cos\left[\omega_0 t + \alpha \int_0^t m(\lambda)d\lambda\right] + n(t) \tag{7}$$

where $r(t)$=received signal (FM signal + noise)
$m(t)$=message
$n(t)$=additive noise

We might know statistical properties of the message and noise signals. Pdf is usually unknown, but some first and second order moments i.e. mean and autocorrelation are known.

Objective
To find an estimate of the message $\hat{m}(t)$ from $r(t)$ such that

$$E\left[\left|m(t) - \hat{m}(t)\right|^2\right]$$

the mean square error is minimized. This is called the minimum mean squared error (MMSE) estimator.

3. ELEMENTS OF HYPOTHESIS TESTING

Testing a hypothesis is to construct an experiment of chance related to the state of the nature and based on the outcomes of that experiment, decide whether the hypothesis can be accepted or rejected.

Problem formulation

1. Source model: Source is denoted by a set of M hypotheses (possibilities)
   H_i; $i = 1, 2, ..., M - 1$
   with a priori probabilities
   P_i=Prob(source generating H_i)

2. Observation model: This is something similar to impulse response/transfer function of a LTI system. However, for a LTI system, the impulse response is a deterministic function. We use observation (channel) model that are probabilistic.

   - Probabilistic transitions between the source and the observer is given by $p(\vec{R}|H_i)$
   where $\vec{R}_{1\times N}$ is the observation vector and $p(\vec{R}|H_i)$ is the pdf of the observation vector given that the source generated the hypothesis H_i.

3. Decision rule: $D(R)$ is the means by which the observation space called $\mathbb{Z}$, i.e. ($\vec{R} \in \mathbb{Z}$), is decomposed into M disjoint regions $\mathbb{Z}_i \in \mathbb{Z}$
$i = 0, 1, ...M-1$ and

$$\mathbb{Z} = \bigcup_{i=0}^{M-1} \mathbb{Z}_i$$

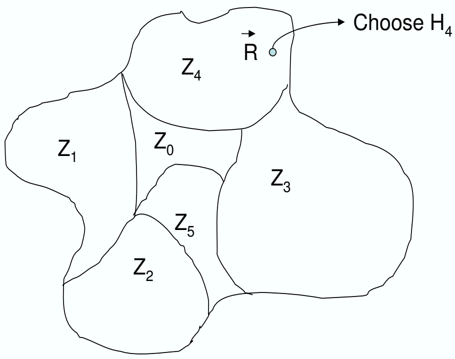

Figure 2. Partitioning of $\mathbb{Z}$

such that H_i is decided iff $R \in \mathbb{Z}_i$.

- $\mathbb{Z}$ is created by the source and observation channel.
- $\mathbb{Z}_i$'s are created by the user based on some reward/punishment criterion. Our job is to select the partition of $\mathbb{Z}$ such that the penalty of wrong decision s minimized.

Example

A source transmits binary data based on this model:

$$S = \begin{cases} 3 & ; \text{under } H_0 \text{ with probability } P_0{=}0.2 \\ -24 & ; \text{under } H_1 \text{ with probability } P_1{=}1-P_0{=}0.8 \end{cases}$$

The received signal is
$R = S + N$
where N is Gaussian noise $\sim N(0, \sigma_n^2)$
Note: $N \in (-\infty, \infty)$
The observation space is $\mathbb{Z} \in (-\infty, \infty)$.

Probability transition (channel pdf):

$$p(R\big|H_0) = \frac{1}{\sqrt{2\pi}\sigma_n} exp\Big[-\frac{(R-3)^2}{2\sigma^2}\Big]$$

since $(R|H_0) \sim N(3, \sigma^2)$ and

$$p(R|H_1) = \frac{1}{\sqrt{2\pi}\sigma_n} exp\Big[-\frac{(R+24)^2}{2\sigma^2} \Big]$$

since $(R|H_1) \sim N(-24, \sigma^2)$.

Binary hypothesis testing

$$H_0 : R = N$$

$$H_1 : R = N + S$$

where S is a random variable with pdf $p_S(s)$.
N is a r.v. with pdf $p_N(n)$.
S and N are independent r.v.'s.

To determine the channel pdf

$$H_0 : R = N; \Rightarrow p(R|H_0) = p_N(R)$$

$$H_1 : R = N + S \Rightarrow p(R|H_1) = p_N(R) * p_S(R) = \int_{-\infty}^{\infty} p_N(R-\lambda) p_S(\lambda) d\lambda$$

Special case: Suppose $S = S_0$; where S_0 is a known constant (fixed).

$$p_S(s) = \delta(s - S_0)$$

or

$$p_S(\lambda) = \delta(\lambda - S_0)$$

substituting this in $p(R|H_1)$:

$$p(R|H_1) = \int_{-\infty}^{\infty} p_N(R-\lambda)\delta(\lambda - S_0) d\lambda$$

$$= p_N(R - S_0)$$

4. BAYES DECISION RULE

Assumptions:

i) $\{P_i\}$; $i = 0, 1, ..., M-1$ apriori probabilities are known.
ii) There are reasonable costs assigned to every decision that are given by

$$C_{ij} = \text{Cost of deciding } H_i \text{ when } H_j \text{ is true}$$

e.g.: Binary data communication

$$H_0 = \text{Transmit } 0$$

$$H_1 = \text{Transmit } 1$$

$$C_{00} = C_{11} = 0 \text{ (right decision)}$$

$$C_{01} = C_{10} = 0 \text{ (wrong decision)}$$

e.g.: Radar/Sonar

$$C_{00} = C_{11} = 0$$
$$C_{01} = 5 \times 10^9 \text{ dollars}$$
$$C_{10} = 2 \times 10^5 \text{ dollars}$$

Bayes' criterion:

Minimize the average cost (risk in choosing $\{\mathbb{Z}_i\}$)

Risk function:

$$\begin{aligned}\mathbb{R} &= E[\text{cost over the probability space of } H_i\text{'s and } \vec{R}] \\ &= \sum_{j=0}^{M-1} \sum_{i=0}^{M-1} P_j C_{ij} P[\text{choosing } H_i \big| H_j \text{ is true}] \\ &= \sum_{j=0}^{M-1} \sum_{i=0}^{M-1} P_j C_{ij} \int_{\mathbb{Z}_i} p(\vec{R}\big|H_j) d\vec{R}\end{aligned}$$

Objective

To identify these $\mathbb{Z}_i$'s such that the risk function $\mathbb{R}(d)$ is minimized.
We rewrite the risk function via:

$$\mathbb{R} = \sum_{j=0}^{M-1} P_j C_{jj} \int_{\mathbb{Z}_j} p(\vec{R}\big|H_j) d\vec{R} + \sum_{j=0}^{M-1} \sum_{i\neq j, i=0}^{M-1} P_j C_{ij} \int_{\mathbb{Z}_i} p(\vec{R}\big|H_j) d\vec{R}$$

We also know that

$$\int_{\mathbb{Z}_j} p(\vec{R}\big|H_j) d\vec{R} = 1 - \sum_{i\neq j} \int_{\mathbb{Z}_i} p(\vec{R}\big|H_j) d\vec{R}$$

substituting this in the expression for $\mathbb{R}$:

$$\begin{aligned}\mathbb{R} &= \sum_{j=0}^{M-1} P_j C_{jj} \left[1 - \sum_{i\neq j} \int_{\mathbb{Z}_i} p(\vec{R}\big|H_j) d\vec{R}\right] + \sum_{j=0}^{M-1} \sum_{i\neq j, i=0}^{M-1} P_j C_{ij} \int_{\mathbb{Z}_i} p(\vec{R}\big|H_j) d\vec{R} \\ &= \sum_{j=0}^{M-1} P_j C_{jj} - \sum_{j=0}^{M-1} P_j C_{jj} \sum_{i\neq j} \int_{\mathbb{Z}_i} p(\vec{R}\big|H_j) d\vec{R} + \sum_{j=0}^{M-1} \sum_{i\neq j, i=0}^{M-1} P_j C_{ij} \int_{\mathbb{Z}_i} p(\vec{R}\big|H_j) d\vec{R} \\ &= \underbrace{\sum_{j=0}^{M-1} P_j C_{jj}}_{I} + \underbrace{\sum_{j=0}^{M-1} \sum_{i\neq j} P_j (C_{ij} - C_{jj}) \int_{\mathbb{Z}_i} p(\vec{R}\big|H_j) d\vec{R}}_{II}\end{aligned}$$

I is invariant in $\mathbb{Z}_i$'s (or decision rule); thus can be omitted.
II is the one that has to be minimized. In particular, we wish to minimize $\sum_i \int_{\mathbb{Z}_i} I_i(\vec{R}) d\vec{R}$, where

$$I_i(\vec{R}) = \sum_j P_j (C_{ij} - C_{jj}) p(\vec{R}\big|H_j) d\vec{R}$$

We know that:

$$P_j \geq 0 \tag{8}$$

$$p(\vec{R}|H_j) \geq 0 \tag{9}$$

Moreover in any reasonable decision rule, the cost of making a wrong decision C_{ij}, where $i \neq j$, is always greater than the cost of making the right decision C_{jj}. Thus:

$$C_{ij} - C_{jj} \geq 0 \tag{10}$$

which implies that the integrand

$$I_i(\vec{R}) \geq 0 \tag{11}$$

Conclusion: To minimize $\mathbb{R}$, it is sufficient to minimize $I_i(\vec{R})$.

This leads to the following decision rule known as the Bayes Decision rule: Choose $\mathbb{Z}_i$ as the region over which

$$I_i(\vec{R}) = \sum_j P_j(C_{ij} - C_{jj})p(\vec{R}|H_j)$$

$$= \sum_{j \neq i} P_j(C_{ij} - C_{jj})p(\vec{R}|H_j)$$

is the smallest. i.e. decide H_i if $I_i(\vec{R}) < I_k(\vec{R})$.

Procedure
Construct $I_0(\vec{R})$, $I_1(\vec{R})$, ..., $I_{M-1}(\vec{R})$.
Choose the one that is the smallest as the decision.
Assign $\vec{R}$ to $\mathbb{Z}_i$.

Another way to identify Bayes decison rule is:

$$I_k(\vec{R}) \gtrless^{\text{not } H_k}_{\text{not } H_m} I_m(\vec{R})$$

where $k \neq m$ and $k, m = 0, 1, ..., M-1$.