Introduction to Detection Theory

Aleksandar Dogandžić

April 4, 2017

Contents

READING: [Hero 2015, §7] and [Johnson 2013, §5].

Motivation

WE wish to make a decision on a signal of interest using noisy measurements. Statistical tools enable systematic solutions and optimal design.

Application areas include:

- communications,
- radar and sonar,
- nondestructive evaluation (NDE),
- biomedicine, etc.

Example: Radar detection

[Hero 2015, Ex. 31]

WE wish to decide on the presence or absence of a target.

We collect a continuous-time measurement $x(t)$ over an interval $[0, T]$ and wish to decide whether $x(t)$ is only noise

$$x(t) = w(t), \qquad 0 \le t \le T$$

or

$$x(t) = s(t - \tau) + w(t), \qquad 0 \le t \le T$$

where

- $s(t)$ is a known signal,

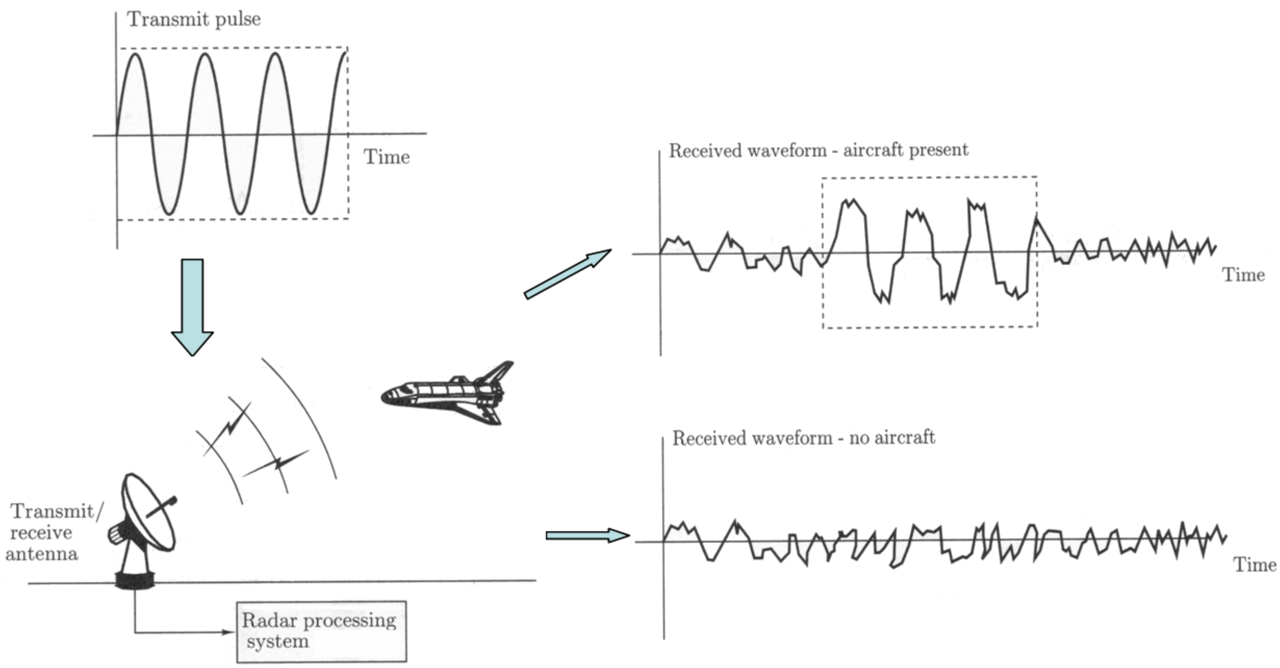

Figure 1: Radar system.

- $w(t)$ is a zero-mean additive white Gaussian noise (AWGN) with known power spectral density (psd) level $\mathcal{N}_0/2$,
- τ is a known time delay, $0 \le \tau \le T$,
- $s(t-\tau)$ is a short-duration "pulse" completely covered by the time interval $[0, T]$; hence,

$$\int_0^T s^2(t-\tau)\,\mathrm{d}t = \int_0^T s^2(t)\,\mathrm{d}t$$

is the signal energy.

⊛ GOAL: Decide reliably if the signal is present or not.
Common notation that has been developed for stating the detection

$$\begin{aligned} \mathbb{H}_0 : &\quad \text{signal absent} \\ \text{versus} & \\ \mathbb{H}_1 : &\quad \text{signal present} \end{aligned}$$

or, in our example,

$$\begin{aligned} \mathbb{H}_0 : &\quad \theta = 0 \\ \mathbb{H}_1 : &\quad \theta = 1 \end{aligned}$$

for the measurement model

$$x(t) = \theta s(t-\tau) + w(t), \qquad 0 \le t \le T.$$

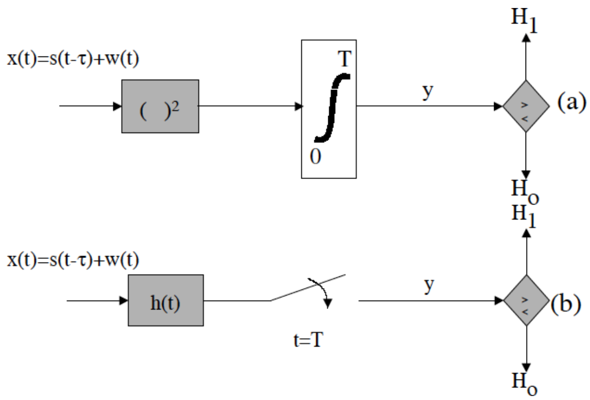

Figure 2: (a) energy and (b) filter detector.

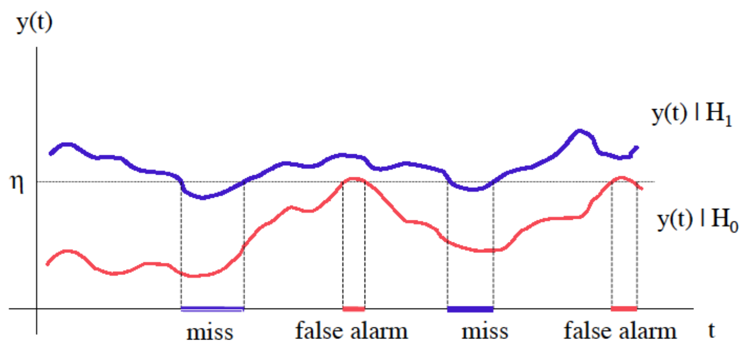

Figure 3: Illustration of probabilities of miss and false alarm.

Consider the two detectors depicted in Fig. 2.

✻ Energy detector:

$$y = \int_0^T x^2(t)\,\mathrm{d}t \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \eta. \tag{1}$$

η is a threshold

✻ Filter detector:

$$y = \underbrace{\int_0^T x(t)h(T-t)\,\mathrm{d}t}_{(x*h)(T)} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \eta. \tag{2}$$

$*$ stands for convolution

Two types of decision errors:

signal absent: false alarm $y > \mu$,

signal present: miss $y < \mu$

see Fig. 3. The probabilities of these errors are computed as follows:

$$P_{\mathrm{FA}} = \Pr\Big\{\text{say signal} \,\big|\, \text{no signal}\Big\} = \int_{y>\eta} f(y \,|\, \mathbb{H}_0)\,\mathrm{d}y$$
$$P_{\mathrm{M}} = \Pr\Big\{\text{say no signal} \,\big|\, \text{signal}\Big\} = \int_{y<\eta} f(y \,|\, \mathbb{H}_1)\,\mathrm{d}y$$

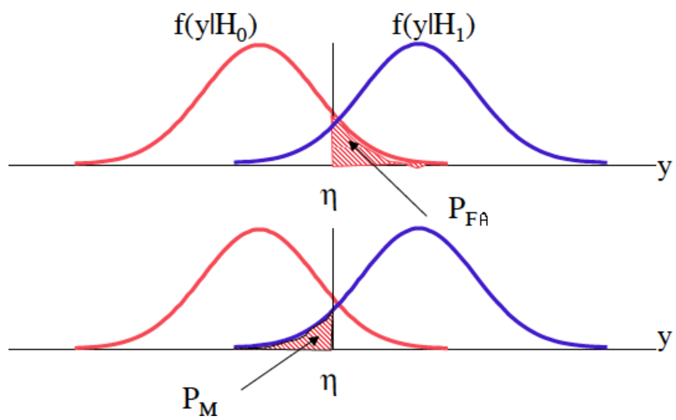

Figure 4: Probabilities of false alarm and miss: decreasing one type of error by changing the decision threshold η necessarily leads to increasing the other type of error.

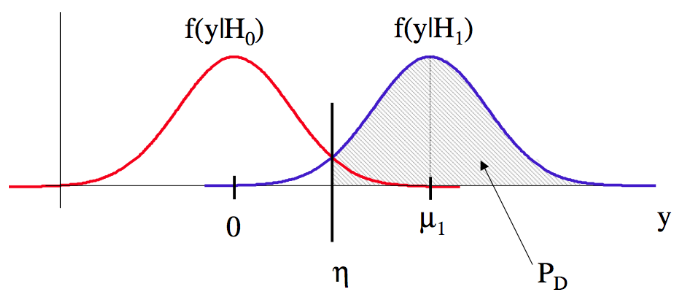

Figure 5: Probability of detection.

where $f(y \mid \mathbb{H}_1)$ and $f(y \mid \mathbb{H}_0)$ denote the distributions[1] of the decision variable y when the signal is present and absent, respectively. See Fig. 4.

☞ An important task in detection theory: find a tradeoff for these two error probabilities, see Fig. 4.

For the filter-threshold detector, we need to design the matched filter $h(t)$: A good strategy is to minimize overlap between the two densities in Fig. 4, measured by the *deflection coefficient* or signal-to-noise ratio (SNR)

$$d^2 = \frac{\left[\mathrm{E}(Y|\mathbb{H}_1) - \mathrm{E}(Y|\mathbb{H}_0)\right]^2}{\mathrm{var}(Y \mid \mathbb{H}_0)}$$

where maximizing d^2 corresponds to minimizing the overlap.

We now compute the deflection for the filter-threshold detector under the above radar scenario. Note that the presence of the signal produces shift in mean but the variance remains the same: $\mathrm{var}\big(Y \mid \mathbb{H}_0\big) = \mathrm{var}\big(Y \mid \mathbb{H}_1\big)$, so the denominator in (3) is the variance of Y under both hypotheses.

$$\begin{aligned} \mathrm{E}(Y \mid \mathbb{H}_0) &= 0 \\ \mathrm{E}(Y \mid \mathbb{H}_1) &= \int_0^T s(t-\tau)h(T-t)\,\mathrm{d}t \\ \mathrm{var}(Y \mid \mathbb{H}_0) &= \mathrm{E}(Y^2 \mid \mathbb{H}_0) \end{aligned}$$

[1] We choose probability density functions (pdfs) here, extends directly to probability mass functions (pmfs).

$\mathrm{E}\big[W(t)W(q)\big] = \frac{\mathcal{N}_0}{2}\delta(t-q)$

$$
\begin{aligned}
&= \mathrm{E}\Big[\int_0^T W(t)h(T-t)\,\mathrm{d}t \int_0^T W(q)h(T-q)\,\mathrm{d}q\Big] \\
&= \int_0^T \int_0^T \mathrm{E}\big[W(t)W(q)\big]h(T-t)h(T-q)\,\mathrm{d}t\,\mathrm{d}q \\
&= \frac{\mathcal{N}_0}{2}\int_0^T h^2(T-t)\,\mathrm{d}t.
\end{aligned}
$$

By the Cauchy-Schwarz inequality,

$$
\begin{aligned}
d^2 &= \frac{2}{\mathcal{N}_0}\frac{\left[\int_0^T s(t-\tau)h(T-t)\,\mathrm{d}t\right]^2}{\int_0^T h^2(T-t)\,\mathrm{d}t} \\
&\le \frac{2}{\mathcal{N}_0}\int_0^T s^2(t-\tau)\,\mathrm{d}t \\
&= \frac{2}{\mathcal{N}_0}\int_0^T s^2(t)\,\mathrm{d}t
\end{aligned}
$$

We assume that the signal $s(t-\tau)$ is entirely inside the time range $t \in (0,T)$.

where the equality holds if and only if

$$
h(T-t) = as(t-\tau), \qquad t \in (0,T)
$$

define $s_1(t) = s(t-\tau)$

for some nonzero constant a; pick $a = 1$ and the SNR-optimal filter is

$$
h(t) = s(T-\tau-t).
$$

The optimal detector can be implemented as:

$$
\begin{aligned}
\int_0^T x(t)h(T-t)\,\mathrm{d}t &= \int_0^T x(t)s(t-\tau)\,\mathrm{d}t \\
&= (x*h)(T) \\
&= x(t)*s(-t)\,|_{t=\tau}
\end{aligned}
$$

Figs. 6 and 7 show implementations of the above SNR-optimal receiver.

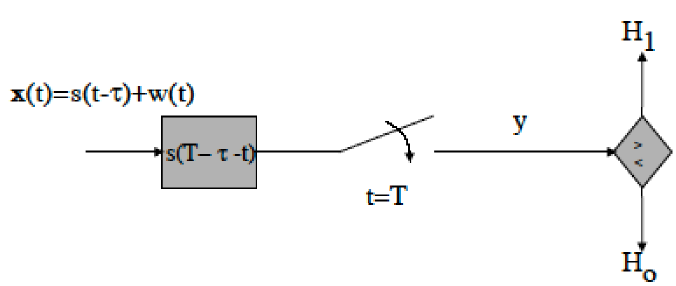

Figure 6: SNR-optimal receiver implemented as a matched filter receiver for delayed signal in noise.

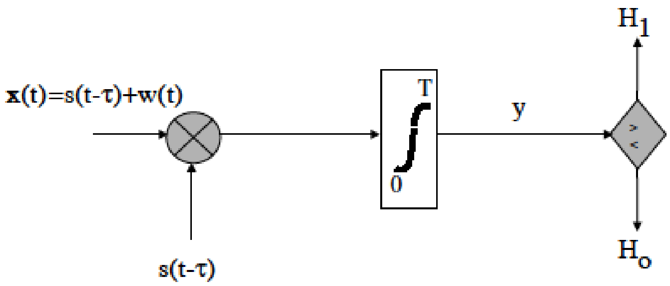

Figure 7: SNR-optimal receiver implemented as a correlator reciever for delayed signal in noise.

General Detection Problem

WE ASSUME a parametric measurement model $f_{X|\Theta}(x|\theta)$.

In point estimation theory, we estimated the parameter $\Theta \in \mathrm{sp}_\Theta$ using the measurement x, where sp_Θ denotes the parameter space for the parameter Θ.

Suppose now that we choose parameter subsets $\mathrm{sp}_\Theta(0)$ and $\mathrm{sp}_\Theta(1)$ that form a *partition* of the parameter space sp_Θ:

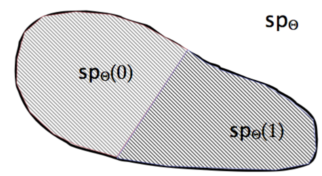

Figure 8: Parameter space partitioning.

$$\mathrm{sp}_\Theta(0) \cup \mathrm{sp}_\Theta(1) = \mathrm{sp}_\Theta, \qquad \mathrm{sp}_\Theta(0) \cap \mathrm{sp}_\Theta(1) = \emptyset$$

In detection theory, we wish to identify *which* hypothesis is true, i.e., *decide*

$$\begin{aligned} \mathbb{H}_0 : &\quad \Theta \in \mathrm{sp}_\Theta(0) \quad \text{null hypothesis versus} \\ \mathbb{H}_1 : &\quad \Theta \in \mathrm{sp}_\Theta(1) \quad \text{alternative hypothesis.} \end{aligned}$$

TERMINOLOGY. If Θ can only take two values,

$$\mathrm{sp}_\Theta = \{\theta_0, \theta_1\}, \quad \mathrm{sp}_\Theta(0) = \{\theta_0\}, \quad \mathrm{sp}_\Theta(1) = \{\theta_1\}$$

we say that the hypotheses are *simple*. Otherwise, we say that they are *composite*.

EXAMPLES. $\mathbb{H}_0 : \Theta = 0$ versus $\mathbb{H}_1 : \Theta \in (0, +\infty)$; here, $\mathrm{sp}_\Theta(0) = \{0\}$ and $\mathrm{sp}_\Theta(1) = (0, +\infty)$.

$\mathbb{H}_0 : \Theta \in (-\infty, 0]$ versus $\mathbb{H}_1 : \Theta \in (0, +\infty)$; here, $\mathrm{sp}_\Theta(0) = (-\infty, 0]$ and $\mathrm{sp}_\Theta(1) = (0, +\infty)$.

DECISION RULE. A decision rule $\phi(\boldsymbol{x})$ maps the *measurement (data) space $\mathcal{X}$* to $\{0, 1\}$:

measurement space $\mathcal{X}$ is the set of values that measurements $\boldsymbol{x}$ can take with nonzero probability

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \text{decide } \mathbb{H}_1, \\ 0, & \text{decide } \mathbb{H}_0 \end{cases}.$$

Here, $\phi(\boldsymbol{x})$ partitions the measurement space $\mathcal{X}$ into two regions:

$$\mathcal{X}_0 = \{\boldsymbol{x} : \phi(\boldsymbol{x}) = 0\} \quad \text{and} \quad \mathcal{X}_1 = \{\boldsymbol{x} : \phi(\boldsymbol{x}) = 1\}.$$

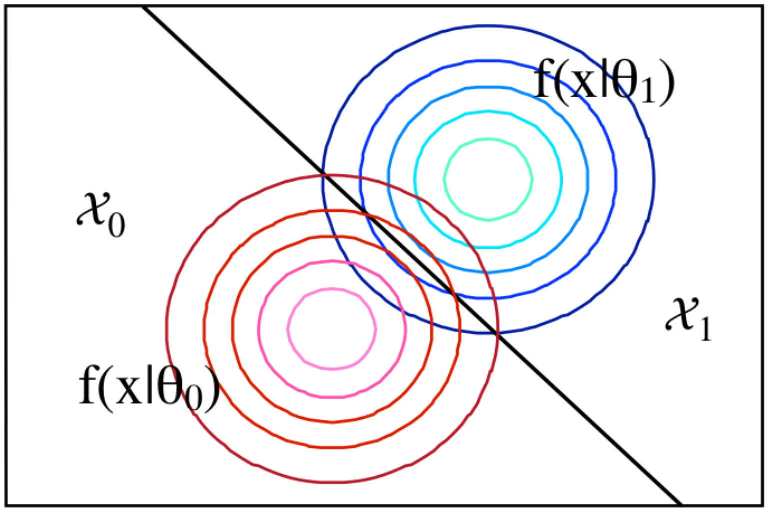

Figure 9: Illustration of decision regions $\mathcal{X}_0$ and $\mathcal{X}_1$ for deciding $\mathbb{H}_0$ and $\mathbb{H}_1$ for an observation x in the plane. Also shown are constant contours of the densities $f_{X\,|\,\Theta}(x\,|\,\theta_0)$ and $f_{X\,|\,\Theta}(x\,|\,\theta_1)$ for $\theta_0 \in \mathrm{sp}_\Theta(0)$ and $\theta_1 \in \mathrm{sp}_\Theta(1)$.

Define the probabilities of false alarm and detection (correctly deciding $\mathbb{H}_1$) for the decision rule $\phi(\boldsymbol{x})$:

$$\begin{aligned}
\mathrm{E}_{\boldsymbol{X}\,|\,\Theta}\big[\phi(\boldsymbol{X})\,|\,\theta\big] &= \int_{\mathcal{X}} \phi(\boldsymbol{x})\, f_{\boldsymbol{X}\,|\,\Theta}(\boldsymbol{x}\,|\,\theta)\,\mathrm{d}\boldsymbol{x} \\
&= \int_{\mathcal{X}_1} f_{\boldsymbol{X}\,|\,\Theta}(\boldsymbol{x}\mid\theta)\,\mathrm{d}\boldsymbol{x} \\
&\triangleq \begin{cases} P_{\mathrm{FA}}(\theta), & \theta \in \mathrm{sp}_\Theta(0) \\ P_{\mathrm{D}}(\theta), & \theta \in \mathrm{sp}_\Theta(1) \end{cases}.
\end{aligned}$$

Then, the probability of miss is

$$\begin{aligned}
P_{\mathrm{M}}(\theta) &= 1 - P_{\mathrm{D}}(\theta) \\
&= \mathrm{E}_{\boldsymbol{X}\,|\,\Theta}\big[1 - \phi(\boldsymbol{x})\,|\,\theta\big] \\
&= \int_{\mathcal{X}_0} f_{\boldsymbol{X}\,|\,\Theta}(\boldsymbol{x}\mid\theta)\,\mathrm{d}\boldsymbol{x}
\end{aligned}$$

for θ in $\mathrm{sp}_\Theta(1)$

see Fig. 5.

☞ STATISTICAL terminology. Statisticians use the following terminology:

- False alarm is "Type I error"
- Miss is "Type II error"
- Probability of detection is "Power"
- Probability of false alarm is "Significance level."

References

Hero, Alfred O. (2015). *Statistical Methods for Signal Processing*. Lecture notes. Univ. Michigan, Ann Arbor, MI (cit. on p. 1).

Johnson, Don H. (2013). *Statistical Signal Processing*. Lecture notes. Rice Univ., Houston, TX (cit. on p. 1).