14: Detector Performance Analysis Techniques

25 feb 2016

In the previous lecture, quadratic detector performance analysis was done. In this lecture, we will look at some detector performance analysis techniques. we will see Chernoff Bound technique and look upon Markov's inequality and Jensen's inequality.

1 Detector performance analysis techniques:

Let us consider likelihood ratio tests (LRTs) for the following example of a composite hypothesis testing problem.

$$\tilde{\delta}(\bar{y}) = \begin{cases} 1, & if\ T(\bar{y}) > \tau', \\ \gamma, & if\ T(\bar{y}) = \tau', \\ 0, & if\ T(\bar{y}) < \tau', \end{cases} \tag{1}$$

where $T : \Gamma \to \mathbb{R}$ is any function. Detection performance typically is a function of two terms,

$$P_F(\tilde{\delta}_T) = P_0[T(y) > \tau] + \gamma \cdot P_0[T(y) = \tau] \tag{2}$$

$$\begin{aligned} P_M(\tilde{\delta}_T) &= P_1[T(y) < \tau] + (1-\gamma) \cdot P_1[T(y) = \tau] \\ &= 1 - P_D(\tilde{\delta}_T) \end{aligned} \tag{3}$$

Where $P_F(\tilde{\delta_T})$ is probability of false alarm and $P_M(\tilde{\delta_T})$ is the probability of miss. In general, if Y has a pdf P_j under H_j,$j \in 0, 1$, then,

$$P_F(\tilde{\delta_T}) = \int \cdots \int_{\{y:T(y)>\tau\} \leq \Gamma(\gamma)} (P_0(y_1 \cdots y_n) \cdot dy_1 \cdots dy_n), \tag{4}$$

which is typically impossible to exactly compute in closed form. As the dimension of Y increases, this integral becomes harder.

1.1 Chernoff Bound technique:

Markov's inequality: If $X \geq 0$ is a random variable,r,then

$$P[X \geq a] \leq \frac{\mathbb{E}[X]}{a}, \forall a > 0. \tag{5}$$

This helps to map probability calculations to expectation calculations which are easier than former.

Jensen's inequality: If X is a random variable and $f : \mathbb{R} \to \mathbb{R}$ is a convex function, then

$$f(\mathbb{E}) \leq \mathbf{E}[f(x)]. \tag{6}$$

Definition 1.1 (Convex function).

$$f[\lambda x + (1-\lambda) \cdot y] \ \leq \ \lambda f(x) + (1-\lambda) \cdot f(y), \ \forall x, y \in \mathbb{R}, \ \forall \lambda \in [0,1]. \tag{7}$$

Now consider,

$$P_F(\tilde{\delta}_T) \ \leq \ P_0[T(Y) \geq \tau]. \tag{8}$$

RHS can be written as,

$$P_0[sT(Y) \geq s\tau] = P_0[e^{sT(Y)} \geq e^{s\tau}], \forall s \geq 0, \tag{9}$$

now both $e^{sT(Y)}$ and $e^{s\tau}$ are positive, so Markov's inequality can be applied. hence,

$$P_0[e^{sT(Y)} \geq e^{s\tau}] \leq e^{-s\tau} \cdot \mathbb{E}_0[e^{sT(Y)}], \forall s \geq 0. \tag{10}$$

RHS can be written as,

$$e^{-s\tau + \log(\mathbb{E}_0[e^{sT(Y)}])} = e^{-s\tau + \mu_{T,0}(S)}, \ \forall s \geq 0. \tag{11}$$

$\log(\mathbb{E}_0[e^{sT(Y)}])$ is the log MGF of $T(Y)$ (Moment Generating Function of T at $s > 0$). Similarly, for probability of miss,

$$\begin{aligned} P_M(\tilde{\delta}_T) &\leq P_1[T(Y) \leq \tau], \\ P_M(\tilde{\delta}_T) &\leq e^{-t\tau + \mu_{T,1}(t)}, \ \forall t < 0. \end{aligned} \tag{12}$$

Hence, we can choose $s > 0, \ t < 0$ to minimize the RHS in eqns. (10) and (12) to obtain the best bound,provided MGFs of T are known.

Lets look at LRTs,where

$$\begin{aligned} T(y) &= \log(L(y)), \\ &= \log\left(\frac{P_1(y)}{P_0(y)}\right), \end{aligned}$$

here, we are comparing the above with a threshold τ, therefore the performance of a detector is a function of τ and $T(y)$. Now,from the definition of $\mu_{T,0}(s)$ and expression for $\mathbb{E}[.]$,

$$\begin{aligned}\mu_{T,0}(s) &= \log\left(\int_\Gamma e^{s\log(L(y))}\cdot P_0(y)\cdot dy\right)\\ &= \log\left(\int_\Gamma (L(y))^s P_0(y)\cdot dy\right)\end{aligned}$$

$$\begin{aligned}\mu_{T,1}(t) &= \log\left(\int_\Gamma (L(y))^t\frac{P_1(y)}{P_0(y)}\cdot P_0(y)\cdot dy\right)\\ &= \log\left(\int_\Gamma (L(y))^{(t+1)}\cdot P_0(y)\cdot dy\right)\\ &= \mu_{T,0}(t+1)\end{aligned}$$

We are trying to bound it such that it increases by s from one side and decreases by $(t+1)$ from other (as t is negative). Therefore, we have the bounds as,

$$\begin{aligned}P_F(\tilde{\delta}_T) &\leq e^{-s\tau+\mu_{T,0}(s)},\ \forall s\geq 0,\\ P_M(\tilde{\delta}_T) &\leq e^{(1-s)\cdot\tau+\mu_{T,0}(s)},\ \forall s<1.\end{aligned}$$

We get the best bound when RHS of both are minimum. RHS of both the terms are similar, except for e^τ term in second expression.

Fact 1. $f(s) = \mu_{T,0}(s) - s\tau$ is a convex function over s. This implies that, if $\mu'_{T,0}(s_0) = \tau,\ \forall s_0$,then f has a minimum at s_0, i.e.,$f(s)$ has its derivative zero at s_0. So, if s_0 lies between 0 and 1, our problem of finding optimum solution will be solved, as it minimizes both the equations. Hence, s_0 is a global minimum and not local minimum.

Fact 2. $\mu'_{T,0}(j) = \mathbb{E}_j[\log(L(y))],\ \forall j\in\{0,1\}$.

Fact 3. $f(s) = \mu_{T,0}(s) - s\tau$ is convex, implies that $\mu'_{T,0}(s)$ is a non decreasing function of s.

Let,

$$\begin{aligned}\mu_0 &= \mu'_{T,0}(0),\\ &= \mathbb{E}_0[\log(L(y))],\\ &= \mathbb{E}_0\left[\log\left(\frac{P_1(y)}{P_0(y)}\right)\right].\end{aligned}$$

Now, using Jensen's inequality, since log is a concave function, we write,

$$\mathbb{E}_0 \left[\log \left(\frac{P_1(y)}{P_0(y)} \right) \right] \leq \log \left(\mathbb{E}_0 \left[\frac{P_1(y)}{P_0(y)} \right] \right). \tag{13}$$

RHS can be written as,

$$\log \left(\int \frac{P_1(y)}{P_0(y)} \cdot P_0(y) \cdot dy \right) = \log(1) = 0. \tag{14}$$

This implies that,

$$\mu'_{T,0}(0) \leq 0. \tag{15}$$

Similarly,

$$\mu'_{T,0}(1) \geq 0. \tag{16}$$

Therefore between 0 and 1, it should cross $\mu'_{T,0}(1) = s$, for a given s.

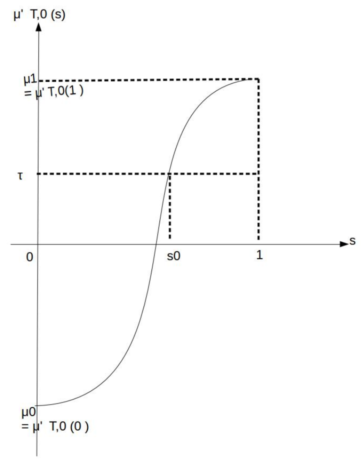

Figure 1: variation of derivative of CGF of $T(y)$ with s. At s_0, τ is attained. *(Source: H. Vincent Poor, An Introduction to Signal Detection and Estimation(Second Edition))*

Remark 1. If $\tau \in (\mu_0, \mu_1)$ then, $\exists s_0 \in [0, 1] : f'(s_0) = 0$ {KL divergence is the dual of $\log(MGF)$. Therefore, $\mathbb{E}_0 \left[\log \left(\frac{P_1(y)}{P_0(y)} \right) \right]$ gives negative KL divergence.}.

We have,

$$P_F(\tilde{\delta}_T) \leq e^{\mu_{T,0}(s_0) - s_0 \mu\prime_{T,0}(s_0)}, \tag{17}$$

$$P_M(\tilde{\delta}_T) \leq e^{\mu_{T,0}(s_0) + (1-s_0)\mu\prime_{T,0}(s_0)}. \tag{18}$$

This is called Chernoff bound technique.

Note 1. When $\tau \leq \mu_0$, $\mu'_{T,0}(s) \geq \tau, \forall s \geq 0$, hence $f'(s) \geq 0$. This means that $\min_{\forall s \geq 0} f(s) = f(0)$, and hence $P_F(\tilde{\delta}_T) \leq 1$.

Note 2. Similar to above, when $\tau \geq \mu_1$, we get, $P_M(\tilde{\delta}_T) \leq 1$.

Example 1.2. For Bayesian hypothesis testing:

Assume prior probabilities are π_0 and π_1 (for H_0 and H_1), then average probability of error is,

$$P_e = \pi_0 P_F + \pi_1 P_M,$$

$$P_e \leq \pi_0 e^{\mu_{T,0}(s_0) - s_0 \mu'_{T,0}(s_0)} + \pi_1 e^{\mu_{T,0}(s_0) - (1-s_0)\mu'_{T,0}(s_0)}. \tag{19}$$

Therefore,

$$P_e \leq [\pi_0 + \pi_1 e^{\mu' T,0(s_0)}] e^{\mu_{T,0}(s_0) - s_0 \mu'_{T,0}(s_0)} \tag{20}$$

(20) is of the form $(A+B)C$ and inequality (19) is obtained from eqns. (17), (18). In fact, one can derive slightly tighter bound than bound obtained in (20),

$$P_e \leq \max\{\pi_0, \pi_1 e^{\mu'_{T,0}(s_0)}\} e^{\mu_{T,0}(s_0) - s_0 \mu'_{T,0}(s_0)}, \ \forall\ 0 \leq s_0 \leq 1 \tag{21}$$

For the detector, $\log(L(y)) \gtreqless \tau$, recalling the minimum probability of error obtained for Bayesian decision rule assuming equal costs, we choose a threshold $\tau = \log\left(\frac{\pi_0}{\pi_1}\right)$. For this detector,

$$P_e \leq \max\{\pi_0, \frac{\pi_0}{\pi_1} \cdot \pi_1\} e^{\mu_{T,0}(s) - s \log\left(\frac{\pi_0}{\pi_1}\right)},$$

$$= \pi_0^{1-s} \cdot \pi_1^s \cdot e^{\mu_{T,0}(s)}, \ \forall\ s \in [0,1].$$

Therefore,

$$P_e \leq= \pi_0^{1-s} \cdot \pi_1^s e^{\mu_{T,0}(s)}, \forall s \in [0,1] \tag{22}$$

Hence, one can try and optimize RHS of above equation over s.

Example 1.3. IID Observations:

For $\Gamma \in \mathbb{R}^n$, and vector $Y = (y_1, y_2, \ldots, y_n) \sim f_j$ under H_j then,

$$
\begin{aligned}
\mu_{T,0}(s) &= \log\left(\mathbb{E}_0\left[e^{s\log(L(y))}\right]\right), \\
&= \log\left(\mathbb{E}_0\left[e^{s\log\left(\prod_{k=1}^{n}\frac{f_1(y_k)}{f_0(y_k)}\right)}\right],\right), \\
&= \log\left(\mathbb{E}_0\left[f_1^s(y)f_0^{-s}(y)\right]\right), \\
&= n\log\left(\int_\Gamma f_1^s(y)f_0^{-s}(y)f_0(y)dy\right), \\
&= n\log\left(\int_\Gamma \left[\frac{f_1(y)}{f_0(y)}\right]^s f_0(y)dy\right).
\end{aligned}
$$

Now,

$$
\int_\Gamma \left[\frac{f_1(y)}{f_0(y)}\right]^s f_0(y)dy = \mathbb{E}\left[\left(\frac{f_1(y)}{f_0(y)}\right)^s\right], \tag{23}
$$

and,

$$
\mathbb{E}\left[\left(\frac{f_1(y)}{f_0(y)}\right)^s\right] \leq \left\{\mathbb{E}_0\left[\frac{f_1(y)}{f_0(y)}\right]\right\}^s \tag{24}
$$

and we know, $\left\{\mathbb{E}_0\left[\frac{f_1(y)}{f_0(y)}\right]\right\}^s = 1$, therefore,

$$
\begin{aligned}
\mathbb{E}\left[\left(\frac{f_1(y)}{f_0(y)}\right)^s\right] &= n\log(a),\ \ a < 1 \\
&= -nc,\ \ c > 0.
\end{aligned}
$$

Note 3. For minimum Probability of error Bayesian detector,

$$
P_e \leq \pi_0^{1-s} \cdot \pi_1^s e^{-nc(s)},\ \ \forall s \in (0,1),
$$

here $c(s)$ should be positive. Probability of making an error decreases exponentially when n increases.

Note 4. By Chernoff's theorem, optimal choice for $s \in (0,1)$ is given by,

$$
\max_{s\in(0,1)} c(s) = D(f||f_0)
$$

where $\{f : D(f||f_0) = D(f||f_1)\}$.