Some Comparison Theorems for Minimax Detection of Gaussian Stochastic Signals

Marat V. Burnashev
Institute for Information Transmission Problems, Moscow, Russia
and Yuri Golubev
CNRS, Université Aix-Marseille, Marseille, France

Le Mans, March 17-20, 2015

1. Simple hypotheses

$\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ – observations

There are two simple hypotheses:

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{y} = \boldsymbol{\xi}, \quad (\text{“only noise''}), \\ \mathcal{H}_1 &: \mathbf{y} = \mathbf{s} + \boldsymbol{\xi}, \quad (\text{“noise } + \text{ stoch. signal''}), \end{aligned} \tag{1}$$

$\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$ – Gaus. rand. vector with distr. $\mathcal{N}(\mathbf{0}, \mathbf{I}_n)$,
$\mathbf{s} = (s_1, \dots, s_n)$ – indep. on $\boldsymbol{\xi}$ Gaus. rand. vector with distr. $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$,
$\boldsymbol{\Sigma} = \mathrm{diag}\{\sigma_i^2 = \mathbf{E}(s_i^2)\}$ –
given $n \times n$-diagonal covariance matrix.

Then

$$p(\mathbf{y}|\mathcal{H}_0) = (2\pi)^{-n/2} e^{-\frac{1}{2}\sum\limits_{i=1}^{n} y_i^2},$$

$$p(\mathbf{y}|\mathcal{H}_1) = \frac{(2\pi)^{-n/2}}{\prod\limits_{1}^{n}(1+\sigma_i^2)^{1/2}} e^{-\frac{1}{2}\sum\limits_{i=1}^{n} y_i^2/(1+\sigma_i^2)}$$

$\Rightarrow$ for likelihood ratio

$$r(\mathbf{y}) = \ln \frac{p(\mathbf{y}|\mathcal{H}_1)}{p(\mathbf{y}|\mathcal{H}_0)} = \frac{1}{2}\sum_{i=1}^{n} \frac{\sigma_i^2 y_i^2}{1+\sigma_i^2} - \frac{1}{2}\sum_{i=1}^{n}(1+\sigma_i^2).$$

Optimal decision (Neyman-Pearson):

$$\sum_{i=1}^{n} \frac{\sigma_i^2 y_i^2}{1+\sigma_i^2} \le A \Rightarrow \mathcal{H}_0, \qquad \sum_{i=1}^{n} \frac{\sigma_i^2 y_i^2}{1+\sigma_i^2} > A \Rightarrow \mathcal{H}_1 \tag{2}$$

Level A is determined by given type I error probablity (false alarm)

$$\alpha = \mathbf{P}(\mathcal{H}_1|\mathcal{H}_0) = \mathbf{P}\left(\sum_{i=1}^{n} \frac{\sigma_i^2 \xi_i^2}{1+\sigma_i^2} > A\right) \tag{3}$$

Type II error probablity (miss probablity) is defined by

$$\beta = \mathbf{P}(\mathcal{H}_0|\mathcal{H}_1) = \mathbf{P}\left(\sum_{i=1}^{n} \frac{\sigma_i^2 y_i^2}{1+\sigma_i^2} \le A\right) = \mathbf{P}\left(\sum_{i=1}^{n} \sigma_i^2 \xi_i^2 \le A\right) \tag{4}$$

For small α, β level A should satisfy constraints

$$\sum_{i=1}^{n} \frac{\sigma_i^2}{1+\sigma_i^2} < A < \sum_{i=1}^{n} \sigma_i^2 \tag{5}$$

Lower bound corresponds to small α (but $\not\to 0$). Upper bound corresponds to small β (but $\not\to 0$).

2. Simple hypothesis against composite altern.

Replace simple altern. $\mathcal{H}_1 = \mathbf{\Sigma} = \mathrm{diag}\{\sigma_i^2\}$ by comp. altern. $\mathcal{H}_1 = \{\mathcal{F}\}$.

$\mathcal{F} = \{\Lambda\}$ – given set of $n \times n$-diagonal covariance matrices.

For decision set $\mathcal{A}$:

$$\begin{aligned} \mathbf{y} \in \mathcal{A} &\Rightarrow \mathcal{H}_0, \\ \mathbf{y} \notin \mathcal{A} &\Rightarrow \mathcal{H}_1, \end{aligned} \tag{6}$$

define error probabilities (minimax statement):

$$\begin{aligned} \alpha = \alpha(\mathcal{A}) = \mathbf{P}(\mathcal{H}_1|\mathcal{H}_0) = \mathbf{P}\left(\mathbf{y} \notin \mathcal{A}|\mathbf{I}_n\right), \\ \beta = \beta(\mathcal{A}, \mathcal{F}) = \sup_{\Lambda \in \mathcal{F}} \mathbf{P}\left(\mathbf{y} \in \mathcal{A}|\mathbf{I}_n + \Lambda\right). \end{aligned} \tag{7}$$

$\Rightarrow$ For given α we want to find

$$\beta(\alpha, \mathcal{F}) = \inf_{\mathcal{A}} \beta(\alpha, \mathcal{A}, \mathcal{F})$$

over all decision sets $\mathcal{A}$.

Clearly, for any α and $\mathcal{F}$

$$\beta(\alpha, \mathcal{F}) \geq \sup_{\Lambda \in \mathcal{F}} \beta(\alpha, \Lambda). \tag{8}$$

Question: For what sets $\mathcal{F}$ we have

$$\beta(\alpha, \mathcal{F}) = \sup_{\Lambda \in \mathcal{F}} \beta(\alpha, \Lambda) \text{ ?} \tag{9}$$

In other words, when it is possible to replace (without performance loss) whole set $\mathcal{F}$ by one its element Λ ?

If $\sup\limits_{\Lambda\in\mathcal{F}} \beta(\alpha,\Lambda)$ is attained for $\Lambda = \Sigma = \mathrm{diag}\{\sigma_i^2\}$, then we need

$$\beta(A,\mathbf{\Lambda}) = \mathbf{P}\left(\sum_{i=1}^{n} a_i^2\xi_i^2 < A\right) \leq \tag{10}$$
$$\leq \beta(A,\mathbf{\Sigma}) = \mathbf{P}\left(\sum_{i=1}^{n} \sigma_i^2\xi_i^2 < A\right),$$

where

$$a_i^2 = \frac{\sigma_i^2(1+\lambda_i^2)}{1+\sigma_i^2} = \sigma_i^2 + \frac{\sigma_i^2(\lambda_i^2-\sigma_i^2)}{1+\sigma_i^2}, \quad i = 1,\ldots,n. \tag{11}$$

In particular, we need

$$\sum_{i=1}^{n} a_i^2 - \sum_{i=1}^{n} \sigma_i^2 = \sum_{i=1}^{n} \frac{\sigma_i^2(\lambda_i^2-\sigma_i^2)}{1+\sigma_i^2} \geq 0. \tag{12}$$

3. Comparison Theorems.

Equivalently, introduce probabilities

$$\beta(x, \boldsymbol{\lambda}) = \mathbf{P}\left(\sum_{i=1}^{n} \lambda_i \xi_i^2 < x\right),$$

$$\beta(x, \boldsymbol{\mu}) = \mathbf{P}\left(\sum_{i=1}^{n} \mu_i \xi_i^2 < x\right),$$

$\xi_1, \ldots, \xi_n$ – indep. $\mathcal{N}(0,1)$-Gaus. rand. variables,
$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ and
$\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n \geq 0$.
When we have (comparison theorem)

$$\beta(x, \boldsymbol{\lambda}) \leq \beta(x, \boldsymbol{\mu}) \ ? \tag{13}$$

We found only one such inequality: Assume that

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0,$$
$$\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n \geq 0,$$

and

$$\sum_{i=1}^{k} \lambda_i \geq \sum_{i=1}^{k} \mu_i, \qquad k = 1, \ldots, n. \tag{14}$$

Then (Bakirov-1995)

$$\beta(x, \boldsymbol{\lambda}) \leq \beta(x, \boldsymbol{\mu}), \tag{15}$$

$$\text{for any} \quad x \geq 2 \sum_{i=1}^{n} \mu_i. \tag{16}$$

Fine inequality (15) is not applicable in our detection problems due to constraint (16) !

We are interested in the case ((5) contradicts to (16))

$$x < \sum_{i=1}^{n} \mu_i. \tag{17}$$

Remark. In order to have (14) fulfilled we need, in particular,

$$\max \lambda_i \geq \max \mu_i,$$

that is rather restrictive.
We present another comparison result.

Proposition 1. *Let*

$$\mu_1 \geq \ldots \geq \mu_n > 0, \qquad b_1 \geq \ldots \geq b_n > 0.$$

If x *satisfies condition*

$$x \max_i \left(\frac{1}{\mu_i} - \frac{1}{b_i} \right)_+ \leq \sum_{i=1}^{n} \ln \frac{b_i}{\mu_i}, \tag{18}$$

then

$$\beta(x, \mathbf{b}) = \mathbf{P} \left(\sum_{i=1}^{n} b_i \xi_i^2 < x \right) \leq \beta(x, \boldsymbol{\mu}).$$

Proof. Denoting

$$\mathcal{A}(x,\boldsymbol{\mu}) = \left\{ \mathbf{y} : \sum_{i=1}^{n} \mu_i y_i^2 \le x \right\},$$

consider difference

$$(2\pi)^{n/2}\left[\beta(x,\mathbf{b}) - \beta(x,\boldsymbol{\mu})\right] = \Delta(x,\mathbf{b},\boldsymbol{\mu}) =$$
$$= \int_{\mathcal{A}(x,\mathbf{b})} \cdots \int e^{-\frac{1}{2}\sum\limits_{i=1}^{n} y_i^2} d\mathbf{y} - \int_{\mathcal{A}(x,\boldsymbol{\mu})} \cdots \int e^{-\frac{1}{2}\sum\limits_{i=1}^{n} y_i^2} d\mathbf{y}.$$

Changing variables $z_i = \sqrt{b_i/\mu_i}\, y_i$, $i = 1,\ldots,n$, and denoting

$$D = \left(\prod_{i=1}^{n} \frac{b_i}{\mu_i} \right)^{1/2}, \qquad d = \max_i \left(\frac{1}{\mu_i} - \frac{1}{b_i} \right)_+ ,$$

we have

$$\Delta(x, \mathbf{b}, \boldsymbol{\mu}) =$$

$$= \int_{\mathcal{A}(x,\mathbf{b})} \cdots \int e^{-\frac{1}{2}\sum\limits_{i=1}^{n} y_i^2} \left[1 - De^{-\frac{1}{2}\sum\limits_{j=1}^{n}\left(\frac{1}{\mu_j} - \frac{1}{b_j}\right) b_j y_j^2} \right] d\mathbf{y} \le$$

$$\le \left(1 - De^{-xd/2}\right) \int_{\mathcal{A}(x,\mathbf{b})} \cdots \int e^{-\frac{1}{2}\sum\limits_{i=1}^{n} y_i^2} d\mathbf{y}.$$

Therefore $\Delta(x, \mathbf{b}, \boldsymbol{\mu}) \le 0$, if $1 - De^{-xd/2} \le 0$, i.e. if condition (18) is satisfied. $\square$

Proposition 1 can be strengthened as follows.
Introduce values $D(k)$, $k = 1, \dots, n$ as unique root of equation

$$\sum_{i=k}^{n} \ln \frac{b_i}{\mu_i} + \sum_{i=k}^{n} \ln[1 - D(k)\mu_i] = 0. \tag{19}$$

Proposition 2. *Let*

$$\mu_1 \geq \dots \geq \mu_n > 0, \qquad b_1 \geq \dots \geq b_n > 0$$

If x satisfies condition

$$x \max_{1 \leq k \leq n} D(k) \leq \sum_{i=1}^{n} \ln \frac{b_i}{\mu_i}, \tag{20}$$

then

$$\beta(x, \mathbf{b}) = \mathbf{P}\left(\sum_{i=1}^{n} b_i \xi_i^2 < x\right) \leq \beta(x, \boldsymbol{\mu}).$$

Example 1. Let $\mu_1 = \ldots = \mu_n = \mu$. Then

$$\max_{1 \le k \le n} D(k) = D(n) = \frac{1}{\mu} - \left(\prod_{i=1}^{n} b_i \right)^{-1/n}$$

and $\beta(x, \mathbf{b}) \le \beta(x, \boldsymbol{\mu})$, if

$$x \le \mu n \qquad \text{and} \qquad \mu^n \le \prod_{i=1}^{n} b_i .$$

Back to signal detection: let
$\mu_i = \sigma_i^2$, $i = 1, \ldots, n$ (see (4) and (11)) and

$$b_i = \frac{\sigma_i^2(1+\lambda_i^2)}{1+\sigma_i^2} = \sigma_i^2 + \frac{\sigma_i^2(\lambda_i^2 - \sigma_i^2)}{1+\sigma_i^2}, \qquad i = 1, \ldots, n.$$

Then from (18) we get
Corollary 1. *If for $n \times n$-diagonal covariance matrix $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_i^2\}$*

$$A \max_i \frac{\lambda_i^2 - \sigma_i^2}{\sigma_i^2(1+\lambda_i^2)} \le \sum_{i=1}^{n} \ln \frac{1+\lambda_i^2}{1+\sigma_i^2}, \tag{21}$$

then $\beta(A, \mathbf{\Lambda}) \le \beta(A, \mathbf{\Sigma})$.

4. Comparison Exponents (Large Deviations).

We find asymptotics of $\beta(A, \boldsymbol{\Sigma})$ and $\beta(A, \boldsymbol{\Lambda})$ as $n \to \infty$. For any $\nu \geq 0$ we have

$$\beta(A, \boldsymbol{\Sigma}) = \mathbf{P}\left(\sum_{i=1}^{n} \sigma_i^2 \xi_i^2 < A\right) \leq$$

$$\leq e^{\nu A/2} \mathbf{E} e^{-\nu \sum\limits_{i=1}^{n} \sigma_i^2 \xi_i^2 / 2} = e^{-g(\nu)},$$

$$2g(\nu) = \sum_{i=1}^{n} \ln(1 + \nu \sigma_i^2) - \nu A,$$

$$2g'(\nu) = \sum_{i=1}^{n} \frac{\sigma_i^2}{1 + \nu \sigma_i^2} - A, \qquad g''(\nu) < 0.$$

If $g'(0) = \sum_{i=1}^{n} \sigma_i^2 - A > 0$, then $\max_{\nu \geq 0} g(\nu)$ is attained for $\nu_0 > 0$, which is defined by equation

$$\sum_{i=1}^{n} \frac{\sigma_i^2}{1 + \nu_0 \sigma_i^2} = A. \tag{22}$$

Then

$$\beta(A, \mathbf{\Sigma}) \leq e^{-g(\nu_0)}. \tag{23}$$

Provided some conditions it is exact asymptotics.

Similarly for any $\mu \geq 0$ we have

$$\begin{aligned} \beta(A, \boldsymbol{\Lambda}) &= \mathbf{P}\left(\sum_{i=1}^{n} a_i^2 \xi_i^2 < A\right) \leq e^{-f(\mu)}, \\ 2f(\mu) &= \sum_{i=1}^{n} \ln(1 + \mu a_i^2) - \mu A, \\ 2f'(\mu) &= \sum_{i=1}^{n} \frac{a_i^2}{1 + \mu a_i^2} - A, \qquad f''(\mu) < 0, \\ a_i^2 &= \frac{\sigma_i^2(1 + \lambda_i^2)}{1 + \sigma_i^2} = \sigma_i^2 + \frac{\sigma_i^2(\lambda_i^2 - \sigma_i^2)}{1 + \sigma_i^2}. \end{aligned} \tag{24}$$

If $f'(0) = \sum_{i=1}^{n} a_i^2 - A > 0$, then $\max_{\mu \geq 0} f(\mu)$ is attained for $\mu_0 > 0$, which is defined by equation

$$\sum_{i=1}^{n} \frac{a_i^2}{1 + \mu_0 a_i^2} = A. \tag{25}$$

Then

$$\beta(A, \mathbf{\Lambda}) \leq e^{-f(\mu_0)}. \tag{26}$$

When $g(\nu_0) \leq f(\mu_0)$?
For example, it is sufficient to have $g(\nu_0) \leq f(\nu_0)$.

Since

$$f(\nu) - g(\nu) = \frac{1}{2} \sum_{i=1}^{n} \ln \frac{1 + \nu a_i^2}{1 + \nu \sigma_i^2} =$$

$$= \frac{1}{2} \sum_{i=1}^{n} \ln \left[1 + \frac{\nu \sigma_i^2 (\lambda_i^2 - \sigma_i^2)}{(1 + \sigma_i^2)(1 + \nu \sigma_i^2)} \right],$$

$\Rightarrow$ it is sufficient to have

$$\sum_{i=1}^{n} \ln \left[1 + \frac{\nu_0 \sigma_i^2 (\lambda_i^2 - \sigma_i^2)}{(1 + \sigma_i^2)(1 + \nu_0 \sigma_i^2)} \right] \geq 0. \qquad (27)$$

If α is small, but $\alpha \not\to 0$, then $\nu_0 \approx 1$.
$\Rightarrow$ sufficient condition for $g(\nu_0) \leq f(\nu_0)$ is

$$\sum_{i=1}^{n} \ln \left[1 + \frac{\sigma_i^2 (\lambda_i^2 - \sigma_i^2)}{(1 + \sigma_i^2)^2} \right] \geq 0. \qquad (28)$$

Estimating $\beta(A, \mathbf{\Sigma})$ from below, we can get
Theorem 1. *If*

$$\sum_{i=1}^{n} \ln \left[1 + \frac{\nu_0\sigma_i^2(\lambda_i^2 - \sigma_i^2)}{(1+\sigma_i^2)(1+\nu_0\sigma_i^2)}\right] \geq 0,$$

then

$$\frac{1}{n}\ln\beta(A, \mathbf{\Lambda}) \leq \frac{1}{n}\ln\beta(A, \mathbf{\Sigma}) + \sqrt{\frac{\delta \ln n}{A}} + \frac{e^2 \ln n}{n}, \tag{29}$$

where $\delta = \max \sigma_k^2 - \min \sigma_k^2$.