MINIMAX ROBUST DETECTION OF A KNOWN SIGNAL IN A GENERAL CLASS OF NOISES

Georgy Shevlyakov and Kiseon Kim

Department of Information and Communications, K-JIST
1 Oryong-dong, Puk-gu, Gwangju 500-712, Korea

ABSTRACT

In practical communication environments, it is frequently observed that the underlying noise pdf is not the Gaussian and may vary in a wide range from short-tailed to heavy-tailed forms. To provide stable and high quality of detection of a known signal, we design the asymptotically minimax (in the Huber sense) minimum distance detection rule under rather general conditions of regularity imposed upon noise pdfs and derive the closed expression for its probability of detection error. In several pdf classes, the least favorable pdfs and corresponding minimax detectors are written down. The minimax robust detectors exhibit robustness of detection in heavy-tailed noises and efficiency in short-tailed ones both in asymptotics and on finite samples.

1. INTRODUCTION

Consider the problem of detection of a known signal $\{\theta_i\}_1^N$ in the additive i.i.d. noise $\{n_i\}_1^N$ with pdf f from a class $\mathcal{F}$. Given $\{x_i\}_1^N$, it is necessary to decide whether the signal $\{\theta_i\}_1^N$ is observed. The problem of detection is set up as the problem of hypotheses testing:

$$H_0 : x_i = n_i \text{ versus } H_1 : x_i = \theta_i + n_i,\ i = 1, \ldots, N. \quad (1)$$

Given a pdf f, the classical theory of hypotheses testing yields various optimal (in the Bayesian, minimax, Neyman-Pearson senses) algorithms for the solution of this problem: all the optimal algorithms are based on the value of the likelihood ratio (LR) statistic $T_N(\mathbf{x}) = \prod_i^N f(x_i - \theta_i)/f(x_i)$ that must be compared with a certain threshold. The differences between the aforementioned approaches result only in the values of a threshold.

In many practical problems of radio-location, acoustics, and communications, noise distributions are only partially known. For instance, it may be known that either the underlying pdf is approximately Gaussian, or there is some information on its behavior in the central zone and on the tails, or an impulsive noise may distort the observed signal, etc. For these detection problems, the well-known Huber minimax approach can be used when the LR statistic is constructed for the least favorable density f^* in a class $\mathcal{F}$, and some robust alternatives to the classical methods have been proposed [3, 5]. Recently, some of these approaches have been extended to more complicated static models of signals under the assumptions of the approximately Gaussian character of noise distributions [10]. Heavy-tailed non-Gaussian noise models with finite and infinite variances both for static and dynamic systems are considered in many works, for example, in [8].

In this paper, we generalize our recent result [6] on minimax detection of a constant signal on the case of minimax detection of a known signal of arbitrary shape. Also, we have interest in the models containing heavy-tailed noise pdfs with large or even with infinite variances as well as short-tailed ones with small variances.

This paper is organized as follows. In Section 2, we give the probability of detection error for the proposed minimum distance detection rule and show that this rule is asymptotically minimax robust in the Huber sense. In Section 3, the particular cases of the minimax detector are considered. In Section 4, concluding remarks are made.

2. MAIN RESULT: THE MINIMAX DETECTOR

In this paper, we deal with the minimum distance detection rule

$$\sum_{i=1}^{N} \rho(x_i) \quad \lessgtr \quad \sum_{i=1}^{N} \rho(x_i - \theta_i), \quad (2)$$

where $\rho(z)$ is a loss function characterizing the assumed form of a distance. This choice of such a simple detection rule is mainly determined by the fact that it allows of the direct use of Huber's minimax theory on M-estimators of location [3]. It can be seen that the choice $\rho(z) = -\log f(z)$ yields the optimal LR test statistic minimizing the Bayesian risk with equal costs and prior probabilities of hypotheses. In this case, it is necessary to know exactly the shape of pdf f to figure out the loss function, and the LR-statistics usually behave poorly under the departures from the assumed pdf model.

This work is supported by IT-Professorship program by IITA in part.

To formulate the result, we need the derivative of a loss function $\psi = \rho'$ called a score function, belonging to a class Ψ.

Assume the following sufficient conditions of regularity imposed upon pdfs f and score functions ψ (they provide consistency and asymptotical normality of M-estimators of location; for details, see [2], pp. 125-127):

($\mathcal{F}1$) f is symmetric and unimodal.

($\mathcal{F}2$) f is twice continuously differentiable on $(0, \infty)$.

($\mathcal{F}3$) Fisher information for location

$$I(f) = \int_{-\infty}^{\infty} [f'(x)/f(x)]^2 f(x)\, dx$$

satisfies $0 < I(f) < \infty$.

($\Psi1$) ψ is well-defined and continuous on $(0, \infty)$.

($\Psi2$) $\int_{-\infty}^{\infty} \psi(x) f(x)\, dx = 0$.

($\Psi3$) $\int_{-\infty}^{\infty} \psi^2(x) f(x)\, dx < \infty$.

($\Psi4$) $0 < \int_{-\infty}^{\infty} \psi'(x) f(x)\, dx < \infty$.

Then, under conditions ($\mathcal{F}1$)–($\Psi4$), the probability of detection error for the minimum distance detection rule (2) takes the following form as $N \to \infty$:

$$P_E = \mathcal{Q}\left(\frac{1}{2}\sqrt{\frac{E}{V(\psi, f)}}\right), \tag{3}$$

where $\mathcal{Q}(z)$ is the complementary error function;
$E = \lim_{N\to\infty} \sum_1^N \theta_i^2$ is the signal energy;

$$V(\psi, f) = \int_{-\infty}^{\infty} \psi^2(x) f(x)\, dx \left[\int_{-\infty}^{\infty} \psi'(x) f(x)\, dx\right]^{-2}$$

is the asymptotic variance of the M-estimators of a location parameter [3];
the signal $\{\theta_i\}_1^N$ is assumed weak in the sense that its amplitudes form the decreasing with N sequences:

$$\theta_i = \theta_{iN} = A_i/\sqrt{N}, \quad i = 1, \ldots, N.$$

From Eq. (3) it directly follows that the minimax problem with respect to the probability of detection error

$$\min_{\psi\in\Psi} \max_{f\in\mathcal{F}} P_E(\psi, f)$$

is equivalent to Huber's problem $\min\limits_{\psi\in\Psi} \max\limits_{f\in\mathcal{F}} V(\psi, f)$.

Thus, all the results on the minimax estimation of location are also applicable in this case: the optimal loss function ρ^* in the minimum distance detector is defined by the maximum likelihood choice for the least favorable pdf f^* minimizing Fisher information for location $I(f)$ over the given class $\mathcal{F}$ [3]

$$\rho^*(x) = -\log f^*(x), \quad \psi^*(x) = -f^{*\prime}(x)/f^*(x),$$

$$f^* = \arg\min_{f\in\mathcal{F}} \int_{-\infty}^{\infty} \left[f'(x)/f(x)\right]^2 f(x)\, dx. \tag{4}$$

Further, the saddle-point pair (ψ^*, f^*) provides the guaranteed upper bound upon the probability of detection error P_E

$$P_E(\psi^*, f) \le P_E(\psi^*, f^*) \quad \text{for all} \quad f \in \mathcal{F}.$$

3. THE PARTICULAR CASES OF THE MINIMAX DETECTOR

Within the minimax approach, the choice of a pdf class $\mathcal{F}$ fully determines all the subsequent stages and qualitative character of the corresponding minimax procedure. Below we enlist the qualitatively different examples of pdf classes with the corresponding least favorable pdfs and minimax detectors.

3.1. Example 1: ε-contaminated Gaussian pdfs

Historically the first class of ε-contaminated standard Gaussian distribution was proposed by Huber [3]

$$\mathcal{F}_H = \{f\colon\ f(x) = (1-\varepsilon)N(x; 0, 1) + \varepsilon h(x), \} \tag{5}$$

where $N(x; 0, 1) = (2\pi)^{-1/2}\exp(-x^2/2)$, $h(x)$ is an arbitrary pdf, and ε ($0 \le \varepsilon < 1$) is a contamination parameter characterizing the fraction of contamination and the level of the uncertainty of information about the shape of an underlying noise pdf. In this case, the least favorable density consists of two parts: the Gaussian in the center and the exponential tails given by

$$f_H^*(x) = \begin{cases} (1-\varepsilon)N(x; 0, 1), & \text{for } |x| \le k, \\ \dfrac{1-\varepsilon}{\sqrt{2\pi}} \exp\left(-k\,|x| + \dfrac{k^2}{2}\right), & \text{for } |x| > k, \end{cases}$$

where the dependence $k = k(\varepsilon)$ is tabulated see [3], p. 87).

The optimal score function has the following limited linear form $\psi_H^*(z) = \max\left[-k, \min(z, k)\right]$. The corresponding minimax robust detector is given by Eq. (2) with $\rho(z) = -\log f_H^*(z)$.

3.2. Example 2: nondegenerate pdfs

In the class $\mathcal{F}_1$ of nondegenerate pdfs (with a bounded density value at the center of symmetry), the least favorable density is known to be the Laplace [9]

$$\mathcal{F}_1 = \{f\colon\ f(0) \ge 1/(2a) > 0\},$$

$$f_1^*(x) = L(x;0,a) = (2a)^{-1}\exp(-|x|/a),$$

here the scale parameter a characterizes the pdf dispersion about the center of symmetry. In this case, we have the sign score function $\psi_1^*(z) = \mathrm{sgn}(x)/a$ and the L_1-norm minimax robust detector given by Eq. (2) with $\rho(z) = |z|$.

3.3. Example 3: pdfs with a bounded variance

In the class $\mathcal{F}_2$ of pdfs with an upper-bounded variance, the least favorable density is the Gaussian [4]

$$\mathcal{F}_2 = \left\{ f\colon\ \sigma^2(f) = \int_{-\infty}^{\infty} x^2 f(x)\,dx \le \overline{\sigma}^2 \right\},$$

$$f_2^*(x) = N(x;0,\overline{\sigma})$$

with the corresponding linear score function $\psi_2^*(z) = z/\overline{\sigma}^2$ and the L_2-norm minimax detector given by Eq. (2) with $\rho(z) = z^2$.

3.4. Example 4: pdfs with a bounded variance and pdf value at the center of symmetry

In the class $\mathcal{F}_{12}$ comprising the restrictions of the two classes $\mathcal{F}_1$ and $\mathcal{F}_2$

$$\mathcal{F}_{12} = \left\{ f\colon f(0) \ge \frac{1}{2a} > 0,\ \sigma^2(f) = \int x^2 f(x)dx \le \overline{\sigma}^2 \right\},$$

the least favorable pdf is of the form [7]:

$$f_{12}^*(x) = \begin{cases} N(x;0,\overline{\sigma}), & \text{for} \quad \overline{\sigma}^2/a^2 < 2/\pi, \\ WH(x;0,\nu,\overline{\sigma}), & \text{for} \quad 2/\pi \le \overline{\sigma}^2/a^2 \le 2, \\ L(x;0,a), & \text{for} \quad \overline{\sigma}^2/a^2 > 2. \end{cases}$$

Here $N(x;0,\overline{\sigma})$ and $L(x;0,a)$ are the Gaussian and Laplace pdfs, respectively;
$WH(x;0,\nu,\overline{\sigma})$ is called the Weber-Hermite pdf given by

$$WH(x;0,\nu,\overline{\sigma}) = \frac{\Gamma(-\nu)\sqrt{2\nu+1+1/S(\nu)}}{\sqrt{2\pi}\,\overline{\sigma}\,S(\nu)} \times \mathcal{D}_\nu^2\left(\frac{|x|}{\overline{\sigma}}\sqrt{2\nu+1+1/S(\nu)}\right). \qquad (6)$$

The shape parameter ν takes its values in $(-\infty, 0]$ and depends on the ratio of parameters $\overline{\sigma}$ and a as follows

$$\frac{\overline{\sigma}}{a} = \frac{\sqrt{2\nu+1+1/S(\nu)}\Gamma^2(-\nu/2)}{\sqrt{2\pi}\,2^{\nu+1}\,S(\nu)\,\Gamma(-\nu)}.$$

Further, $\mathcal{D}_\nu(\cdot)$ are the Weber-Hermite functions or the functions of the parabolic cylinder [1];
$S(\nu) = [\psi(1/2-\nu/2) - \psi(-\nu/2)]/2$, and in this context, $\psi(x) = d\ln\Gamma(x)/dx$ is the digamma function.

The Weber-Hermite pdfs (6) arise as the solution to the Euler-Lagrange equation for the variational problem (4). The Gaussian and Laplace pdfs are the particular cases of Eq. (6) when $\nu = 0$ and $\nu \to -\infty$, respectively.

In this case, the minimax detector is of the form:
• with $\overline{\sigma}^2(f) \le 2a^2/\pi$ or with relatively small variances, it is the minimum L_2-norm distance detector;
• with $\overline{\sigma}^2 > 2a^2$ or with relatively large variances, it is the minimum L_1-norm distance detector;
• with relatively moderate variances, it is given by Eq. (2) with $\rho(z) = -\log f_{12}^*(z)$ but it can be rather effectively approximated by the low-complexity minimum L_{p^*}-norm distance detector with the power $p^* \in (1,2)$ given by

$$p^* = \begin{cases} 5.33 - 7.61x + 3.73x^2, & 2/\pi < x \le 1.35, \\ 2.66 - 1.65x + 0.41x^2, & 1.35 < x < 2, \end{cases}$$

where $x = \overline{\sigma}^2/a^2$.

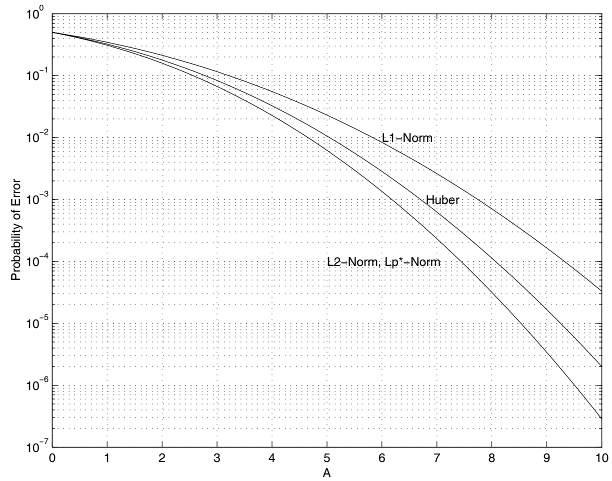

Fig. 1. Gaussian noise: $A = \sqrt{SNR}$.

3.5. Some numerical results on detector performance

The probability of detection error P_E given by Eq. (3) was computed for the standard Gaussian noise, the Cauchy noise with the pdf $f(x) = 1/[\pi(1+x^2)]$, and for the exponential-power noise with the pdf $f(x) \propto \exp(-|x|^q)$.

The performance of the L_1-, L_2-, L_{p^*}-norm, and Huber's detectors was studied. In the latter case, the minimum distance detection rule (2) was examined for Huber's optimal loss function $\rho(z) = -\log f_H^*(z)$ with $\varepsilon = 0.1$. The results of computing are exhibited in Fig. 1 – Fig. 3.

On finite samples with $N = 20$ and $N = 100$, the performance of the minimax detectors was studied by Monte Carlo technique, and the obtained results proved to be close to the asymptotic results obtained from Eq. (3).
Gaussian noise. The L_{p^*}-norm detector coincides with the optimal L_2-norm detector both being better than Huber's.

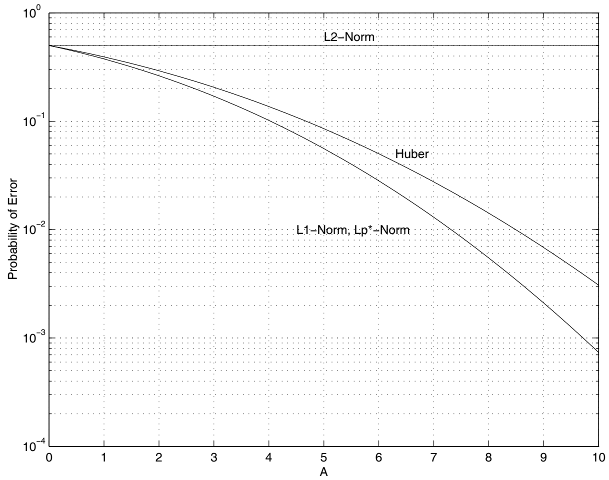

Fig. 2. Cauchy noise: $A = \sqrt{E}$.

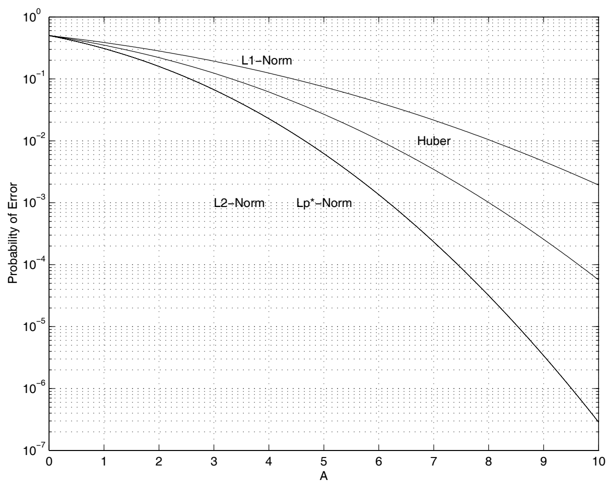

Fig. 3. Exponential-power noise: $q = 100$, $A = \sqrt{SNR}$.

Cauchy noise. The L_2-norm detector naturally has the extremely poor performance. The L_{p^*}-, L_1-norm and Huber's detectors exhibit their good robust properties.
"Short"-tailed noise (the exponential-power with $q = 100$ close to the uniform). The L_2-norm and L_{p^*} detectors prove their superiority over Huber's and the L_1-norm detectors.

4. CONCLUDING REMARKS

Our main aim is to expose some new results on the application of Huber's minimax approach to robust detection:
• First, it is Eq. (3) that yields the asymptotic probability of detection error for the minimum distance detection rule and allows of the direct use of Huber's results on robust estimation of a location parameter in detection problems.
• Second, in short-tailed pdf models it is the significantly better performance of the minimax detector designed for the distribution class with a bounded variance and density value at the center of symmetry as compared to the performance of conventional Huber's soft limited detection rule optimal on the class of ε-contaminated Gaussian distributions. Thus, if the noise variance is actually small enough (as in short-tailed pdf models) then it is quite reasonable to use the L_2-norm detector to enhance the power of detection.
• Moreover, the Gaussian pdf is widely used in applications not because of the CLT arguments which guarantee only the approximate Gaussianity (the main lesson of robust statistics is just that small departures from Gaussianity may cause great consequences [2, 3]) but mainly by the reason that it is the least favorable distribution in models with a bounded variance.
• Finally, within minimax approach it is no need in the detailed specifying of a pdf shape that usually is unrealistic in changing noise environments.

5. REFERENCES

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, New York: Dover, 1972.

[2] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw and W. A. Stahel, *Robust Statistics. The Approach Based on Influence Functions*, New York: Wiley, 1986.

[3] P. J. Huber, *Robust Statistics*, New York: Wiley, 1981.

[4] A. M. Kagan, Yu. V. Linnik, and C. R. Rao, *Characterization Problems in Mathematical Statistics*, New York: Wiley, 1973.

[5] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*, Berlin: Springer, 1988.

[6] G. L. Shevlyakov and Kiseon Kim, "Robust minimax detection of a weak signal in noise with a bounded variance and density value at the center of symmetry", submitted to *IEEE Trans. Inform. Theory*, June 2004.

[7] G. L. Shevlyakov and N. O. Vilchevski, *Robustness in Data Analysis: criteria and methods*, Utrecht: VSP, 2002.

[8] A. Swami and B. Sadler, "On some detection and estimation problems in heavy-tailed noise", *Signal Processing*, vol. 82, pp. 1829-1846, 2002.

[9] B. T. Polyak and Ya. Z. Tsypkin, "Robust identification", In: *Identif. Syst. Parameter Estim., Part 1, Proc. 4th IFAC Symp.*, Tbilisi, pp. 203-224, 1976.

[10] X. Wang and H. V. Poor, "Robust multiuser detection in non-gaussian channels," *IEEE Trans. Signal Proc.*, vol. 47, pp. 289-305, 1999.