MERIDIAN ESTIMATOR PERFORMANCE FOR SAMPLES OF GENERALIZED GAUSSIAN DISTRIBUTION

D.A. Kurkin [1], V.V. Lukin [1], I. Djurovic [2], S. Stankovic [2]

[1] National Aerospace University, 61070, Kharkov, Ukraine
[2] University of Montenegro, 81000, Podgorica, Montenegro

Abstract – **Recently, a sample meridian estimator of location parameter (LP) has been proposed. It has been shown to be robust and adjustable by means of tunable parameter *δ*. Basic properties of this estimator were studied; however its practical behavior needs further analysis. In this paper, we study and analyze the efficiency of the meridian estimator for samples with essentially different properties. For this purpose, we use generalized Gaussian distribution as a model. Also, the efficiency of the meridian estimator is compared to the efficiency of median and myriad estimators.**

I. INTRODUCTION

During many years, Gaussian model has been used widely for noise approximation [1]. However, recent achievements in signal and image processing have shown that in many practical applications noise cannot be adequately approximated by classical Gaussian model. That's why many non-Gaussian models have gained significant popularity. Due to this, it has become necessary to consider robustness aspects while designing methods for data, signal and image processing.

Recently, a new sample meridian estimator has been proposed [2]. It has some similarities to both sample median and myriad estimators. Similarly to the myriad estimator, the sample meridian estimator has a tunable parameter ($δ$) that controls its performance. The sample meridian has been shown to be a quite robust estimator of location parameter (LP), especially for rather small values of tunable parameter $δ$. There are results [3] indirectly indicating that the meridian estimator can be efficient for PDFs with peaky maximum; however, this assumption needs confirmation for a wider set of such distributions. A family of generalized Gaussian distributions (GGD) just represents such properties since it has PDFs with bell-shaped maxima for p>1 and peak-shaped maxima otherwise. Besides, the generalized Gaussian distributions model a wide variety of natural processes and data samples in signal/image processing applications [2].

It is desirable to determine optimal values of tunable parameter $δ$ for different distributions and to compare performance of the meridian estimator to performance of other robust estimators as, e.g., the median and myriad estimators.

II. THEORETICAL BACKGROUND

Let us briefly discuss main properties of GGD. Analytically, generalized Gaussian PDF is defined as

$$f(x) = \frac{p \cdot \alpha}{2\Gamma(1/p)} \exp(\alpha x)^p$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ is gamma function, $\alpha = \sigma_{GGD}^{-1}\sqrt{\Gamma(3/p)/\Gamma(1/p)}$, σ_{GGD} is a standard deviation, and p is a shape parameter. In fact, generalized Gaussian distribution is a parametric family of symmetrical distributions. Distributions with diverse properties can be obtained by varying the shape parameter. The family includes two distributions as special cases: the normal (Gaussian) distribution as a first special case ($p = 2$) and the Laplace distribution ($p = 1$) as a second one.

Analytically a sample meridian estimator of location $\hat{\beta}$ is defined as

$$\hat{\beta} = \arg\min_{\beta}\left[\sum_{i=1}^{N}\ln\left(\delta + |x_i - \beta|\right)\right] = \text{meridian}\{x_i, i = 1, ..., N; \delta\}$$

where N denotes a sample size; x_i is an i–th element of the sample; δ is called medianity parameter.

Let us consider the main properties of the meridian estimate [2].

1) All properties of meridian estimate depend on its error function behavior which is defined as

$$\varphi(\beta) = \sum_{i=1}^{N} \ln\left(\delta + |x_i - \beta|\right);$$

2) The function $\varphi(\beta)$ strictly decreases for $\beta < x_{min}$ and strictly increases for $\beta > x_{max}$ where x_{min} and x_{max} are minimal and maximal elements of an input data sample, respectively.

3) All local minima lie within the range of input sample $[x_{min}, x_{max}]$;

4) As parameter δ increases, properties of the meridian estimator tend to properties of the sample median;

5) The meridian $\hat{\beta}$ is always one of the input samples.

A sample myriad estimator of location $\hat{\gamma}$ is very similar to the sample meridian estimator and is defined as

$$\hat{\gamma} = \arg\min_{\beta}\left[\sum_{i=1}^{N} \ln\left(k^2 + x_i^2\right)\right] = \text{myriad}\{x_i, i = 1, ..., N; \delta\}$$

where k is a myriad estimator tuning parameter. The properties 2) and 3) of the meridian estimator also hold for the myriad estimator. As k increases, properties of the myriad estimator approach to those ones of the mean estimator. The main myriad's difference from meridian is that the myriad $\hat{\gamma}$ is not necessarily one of the input samples. This fact complicates calculation of the myriad estimate.

III. METHODOLOGY OF ANALYSIS

In order to evaluate the performance of the meridian estimator using statistical approach, one should estimate its effectiveness in finding location parameter. Below, this effectiveness is measured with standard deviation σ (SD) and with its robust equivalent – median absolute deviation (MAD) which is defined as

$$MAD = med\{\left|\theta_j - med\{\theta_j\}\right|, j = 1, ..., N\}$$

where θ_j denotes an j-th input sample. For initial data samples, θ_j has to be replaced by x_i (MAD_X), and for the considered LP estimators by either $\hat{\beta}_j$ (MAD_β) or $\hat{\gamma}_j$ (MAD_γ).

Let us determine the probability P that describes the equivalence of the meridian and median estimators: i.e. $meridian\{x_i, i = 1, ..., N; \delta\} = median\{x_i, i = 1, ..., N\}$ when $P = 1$. Fig. 1 demonstrates the estimated relationship between the tuning parameter δ and introduced probability P. We used a meridian estimator with different δ values to estimate LP for samples with 64 elements 1000 times. Afterwards, these dependencies were averaged.
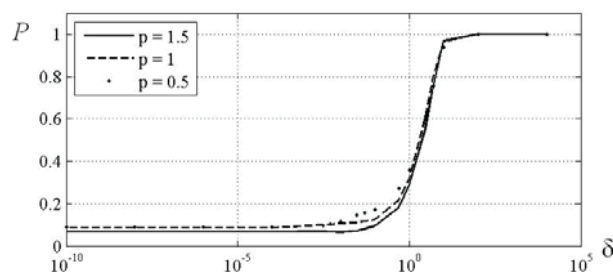

Fig. 1 Estimated relationship between tuning parameter of the meridian estimator and probability of the meridian and median estimators output equivalence.

There are three plots in this Figure and each of them corresponds to a GGD with different values of shape parameter. As it is seen, these dependencies tend to "saturate" to unity if the value of the parameter δ is large enough. Also, they have a flat region if the parameter value is rather small. Transition area between these two zones is observed for δ within the limits from 10^{-2} to 10^{2} or, more generally speaking, for δ/MAD_X from 10^{-2} to 10^{2}. Thus, there is no sense in using meridian estimator with $\delta > 10^{2} \cdot MAD_X$.

IV. ANALYSIS OF OBTAINED RESULTS

As it was said above, the effectiveness of the considered estimators will be evaluated using SD and MAD criteria. Figure 2 demonstrates an example of plots that were obtained to evaluate the effectiveness of the meridian estimator.
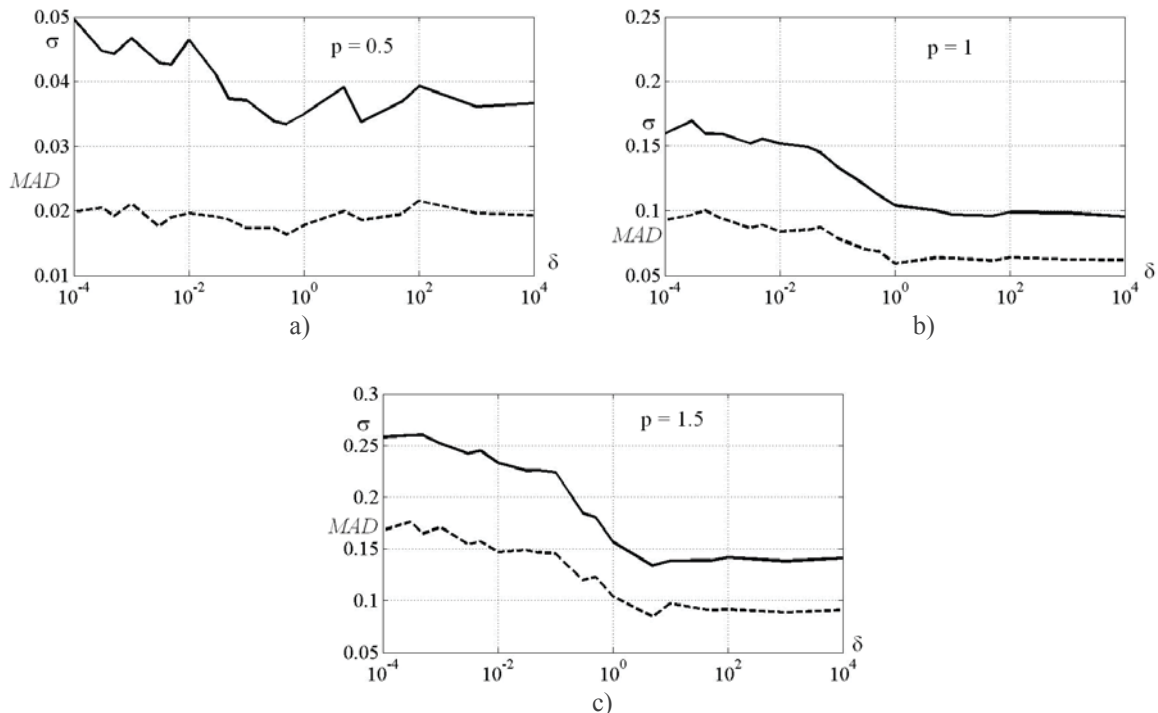

Fig. 2   Dependences of SD and MAD vs δ for meridian estimator for GGD with different values of p: a) p = 0.5; b) p = 1 and c) p = 1.5.

These dependencies were obtained for 1000 realizations for the sample size $N = 64$; $\sigma_{GGD} = 1$. Similar dependences have been also obtained for two other values of N: 128 and 256. The only difference is that the corresponding values of σ and MAD_β are smaller approximately proportionally to $N^{-1/2}$ (see data in Table 1).

Let us analyze these dependencies. For p=1.5, there is a non-obvious minimum observed for δ/MAD_X about 5.0. More obvious minima take place for δ/MAD_X about 1 if p=1 and about 0.5 for p=0.5. More detailed analysis has shown that the meridian estimator produces sufficiently better accuracy than the standard median only for p<0.7.

Also, we should mention that the ratio MAD/σ equals 1.48 for Gaussian distribution. It is clearly seen that for the presented plots $\sigma/MAD_\beta > 1.48$ (Fig. 2). Thus, the distribution law of estimates is not Gaussian and is heavy tailed. In fact, the obtained LP estimates obey a peaky shaped distribution, that's why MAD_β describes the meridian estimator's statistical properties more accurately.

Consider now accuracy of the meridian and myriad estimators. The simulation results are given in Table 1. Subscript opt for SD and MAD means that these are the minimal reachable values. The same subscript for k and δ means that these values provide the best accuracy of LP estimation. Although the values δ_{opt} seem to depend upon the sample size, in fact this is because of wide areas of δ for which MAD_β values are practically constant (see the plot in Fig. 1,a). If p decreases, δ_{opt} reduces as well to provide higher robustness with respect to outliers.

Comparison of MAD_β and MAD_γ shows that these values are practically the same for both considered estimators for p<1 under condition of properly set tuned parameters. For p>1, the myriad estimator is able to provide better accuracy since the meridian estimator in this case performs as the sample median. Thus, the meridian estimator use can be expedient for peaky PDFs. However, it is worth finding a rule for setting δ for a data sample at hand in adaptive manner similarly as it can be done for the myriad estimator [4, 5].

Note that we have carried out studies for other than unity values of σ_{GGD}. They have demonstrated that δ_{opt} is proportional to MAD_X. One more observation is that kurtosis for GGD increases and percentile coefficient of kurtosis [4,5] reduces if p decreases. These can be pre-conditions for design of adaptive meridian estimator applicable in practice.

Table 1 Myriad and meridian estimators effectiveness comparison for GGD with different values of shape parameter

		P	0.3	0.5	0.7	1	1.5
Myriad	N = 64	σ_{opt}	0.00800	0.0347	0.0640	0.0968	0.1251
		MAD_{opt}	0.00266	0.0169	0.0378	0.0642	0.0756
		k_{opt}	10^{-2}	$5\cdot10^{-2}$	$3\cdot10^{-1}$	$7\cdot10^{-1}$	1
	N = 128	σ_{opt}	0.00445	0.0208	0.0439	0.0682	0.0849
		MAD_{opt}	0.00175	0.0118	0.0257	0.0454	0.0565
		k_{opt}	10^{-2}	$3\cdot10^{-2}$	$5\cdot10^{-1}$	$7\cdot10^{-1}$	1
	N = 256	σ_{opt}	0.00260	0.0135	0.0294	0.0467	0.0601
		MAD_{opt}	0.00094	0.0076	0.0176	0.0311	0.0377
		k_{opt}	10^{-3}	$5\cdot10^{-2}$	10^{-2}	$7\cdot10^{-1}$	7
meridian	N = 64	σ_{opt}	0.00815	0.0343	0.0618	0.0971	0.1365
		MAD_{opt}	0.00283	0.0164	0.0350	0.0603	0.0892
		δ_{opt}	10^{-2}	10^{-1}	$5\cdot10^{-1}$	10	10
	N = 128	σ_{opt}	0.00453	0.0208	0.0409	0.0660	0.0934
		MAD_{opt}	0.00168	0.0105	0.0241	0.0399	0.0631
		δ_{opt}	10^{-2}	$5\cdot10^{-1}$	$5\cdot10^{-1}$	3	10
	N = 256	σ_{opt}	0.00248	0.0138	0.0279	0.0454	0.0672
		MAD_{opt}	0.00099	0.0071	0.0170	0.0288	0.0423
		δ_{opt}	10^{-2}	10^{-1}	1	10	10

V. Conclusion

The performed studies have shown that the meridian estimator can be used as an effective estimator for a generalized Gaussian distribution family. Maximum effectiveness can be reached by choosing an optimal value of the tunable parameter δ_{opt}. In this case, the meridian estimator performs either as the sample median or better. The latter is observed if p<1.

The location of δ_{opt} is affected by tail heaviness of an estimated sample distribution and by scale of this distribution (MAD_X). Our future work will be focused on design of adaptive meridian estimator.

References

[1.] A. D. Spaulding, D. Middleton, "Optimum reception in an impulsive interference environment. Part I: Coherent Detection," *IEEE Trans. Commun,* vol. 25, no. 9, pp. 910-923, 1977.
[2.] T. C. Aysal, K. E. Barner, "Meridian Filtering for Robust Signal Processing," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 3949-3962, 2007.
[3.] D. A. Kurkin, V. V. Lukin, A. A. Roenko, I. Djurovic, "Analysis of meridian estimator performance for non-Gaussian PDF data samples," *Telecommunications and Radio Engineering*, vol. 69, no. 8, pp. 669-679, 2010.
[4.] A. A. Roenko, V. V. Lukin, I. Djurovic, "Two approaches to adaptation of sample myriad to characteristics of SαS distribution data," *Signal Processing*, vol. 90, no. 7, pp. 2113-2123, July 2010.
[5.] A. Roenko, V. Lukin, I. Djurovic, "Maximum likelihood, optimal L- and adaptive estimators of location parameter for data samples with symmetric heavy-tailed distributions, " *CD-ROM Proceedings of the 4-th International Symposium on Communications, Control and Signal Procesing, Limassol*, Cyprus, 2010, 4 p.