Title: The Huber's functions and their application to a classification problem

Author: Henryk Gacki, Agnieszka Kulawik

Citation style: Gacki Henryk, Kulawik Agnieszka. (2020). The Huber's functions and their application to a classification problem. "Matematyka Stosowana" (T. 48, no. 2 (2020), s. 101-114), DOI: 10.14708/ma.v48i1.7045

MATHEMATICA APPLICANDA
Vol. 48(2) 2020, p. 101–x
doi: 10.14708/ma.v48i1.7045

Henryk Gacki (Katowice)
Agnieszka Kulawik (Katowice)

The Huber's functions and their application to a classification problem

Abstract In the following paper a classification problem with two multivariate normally distributed classes is considered. The problem is solved in a case of an empirical real situation (a motors data) using the Karhunen-Loeve transform and classifying functions based on estimators for unknown parameters of a multivariate normal distribution. We consider the maximum likelihood estimator and a robust one. The robust estimator bases on the Huber's functions. The corresponding classifying functions (classifiers) are compared using the Leave-One-Out method

2010 Mathematics Subject Classification: Primary: 62C12, Secondary: 62P30.

Key words and phrases: the Huber's function, multivariate normal model, robust estimator, classifier, the Karhunen-Loeve transform.

1. Introduction In the article a classification problem will be considered. The classification is based on empirical discriminant functions for the Gaussian classifier which are defined by estimators of an unknown expected value and a covariance matrix for multivariate normal distributions. In the article several classifiers will be compared for a motors problem. The motors problem has been already used in literature (for example [1], [4]). The problem and the corresponding data were introduced by J. Adamczewski and H. Gacki [1]. In their article the maximum likelihood method was used to solve the classification problem. In the following paper the results will be expanded. The classifiers will base also on a robust estimator and we will try to choose a better estimator for the motors problem then the classical one from [1]. For the robust estimator the Huber's functions play a key role. The function $\phi_t \colon \mathbb{R} \to \mathbb{R}$ was introduced by P. J. Huber in [8], and is defined by its derivative ϕ'_t of form

$$\phi'_t(x) = \begin{cases} x, & |x| \le t, \\ \frac{t^2}{x}, & |x| > t, \end{cases}$$

where $t > 0$. The constant t is called a tunning constant or a truncation level.

The Huber's functions $\phi_t, t > 0$, have a property which is unfavorable for some scientific research. The derivatives ϕ'_t of the Huber's functions are not differentiable. To avoid such undesirable property T. Bednarski and S. Zontek

[2] have presented a modification of the Huber's functions. They propose the modification (see Figure 1) of form

$$\tilde{\phi}_t'(x) = \begin{cases} x, & |x| \le t, \\ -x - 4t - \frac{2t^2}{x}, & -2t < x < -t, \\ -x + 4t - \frac{2t^2}{x}, & t < x < 2t, \\ \frac{2t^2}{x}, & |x| \ge 2t. \end{cases}$$

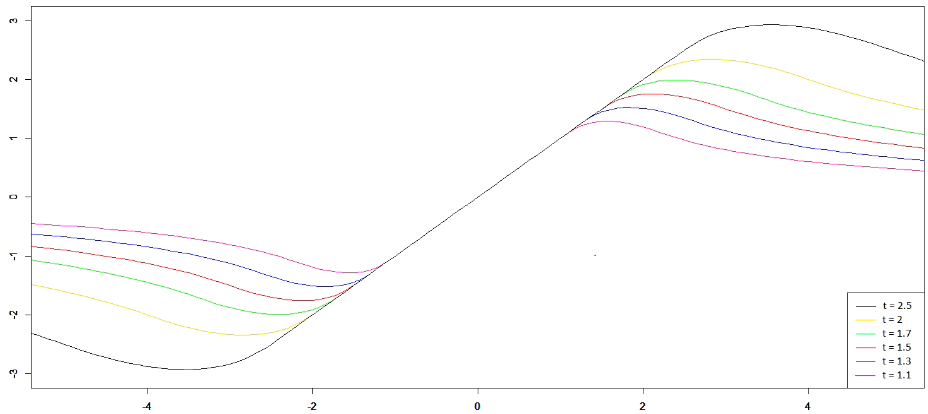

Figure 1: The modifications $\tilde{\phi}_t'$ for $t \in \{1.1, 1.3, 1.5, 1.7, 2, 2.5\}$.

The modifications have been already used in literature. A. Kulawik and S. Zontek [10] used the modification for a robust estimation in the multivariate normal model with variance components. Next they [11] applied the modification in robust estimation of an expected value and a positive definite covariance matrix of a multivariate normal distribution. The modifications were also mentioned by R. Zmyślony and S. Zontek [13].

In the following article we will try to show an opportunity of using the modifications of the Huber's functions in case of an empirical classification problem with two classes. In the next chapter the problem and methods are described precisely. The method bases on the maximum likelihood estimator and a robust estimator for which a Huber's function is needed. The robust estimator is presented in the Chapter 3. In the Chapter 4 the computational results are described. In computations the "R" program has been used. Moreover, the "R" package "expm" (function *sqrtm*) was used in computing the robust estimator (see V. Goulet at al. [6]). The "R" package "conics" (function *conicPlot*) was used to plot the separating surfaces (see B. Desgraupes [3]). The article will end with an open problem about finding the proper Huber's function for the motors problem.

2. A formulation of the problem and assumptions of the experiment. Consider the following situation. Let it be a set whose elements will be called recognition objects. Suppose further that the recognition objects are divided into a finite number of classes, which we will call images. We want to build an algorithm that allows recognition of the object's state class on the basis of previously learned defect class images and to create a criterion for recognizing and making decisions. In our considerations we will assume that there is a certain vector of n measurable features, the determination of which in each of the unpacked objects is always possible, and which allow reliable recognition of the state of the object.

Of the many image recognition methods, it was decided to use ([1]) the Karhunen-Loeve transformation method (the Karhunen-Loeve Transform or shortly - the KLT) in combination with the maximum likelihood method ([5], [9], [12]). The KLT transforms n-dimensional input space Z into an output m-dimensional space Z^* which is a secondary features space, where $m \leq n$.

Let us consider a training set which consists of n-dimensional vectors $\mathbf{y}_j$, $j = 1, \ldots, N$. Assume that each vector belongs to one of M possible pattern classes $\{\omega_i : i = 1, \ldots, M\}$. Let μ_i denote the mean of the random pattern vectors $\mathbf{y}_i$ in the class ω_i, $i = 1, \ldots, M$. Now,

$$\mathbf{z}_i = \mathbf{y}_i - \mu_i$$

denotes the centralized observations from ω_i, $i = 1, \ldots, M$. It is advisable ([5], [9], [12]) to find the covariance matrix $\mathbf{R}$ for the data $\mathbf{z}$:

$$\mathbf{R} = \sum_{i=1}^{M} p_r(\omega_i) E\{\mathbf{z}_i, \mathbf{z}_i^T\},$$

where $p_r(\omega_i)$ denotes a priori probability of the occurrence of the i-th class, $E\{\mathbf{z}_i, \mathbf{z}_i^T\} = \frac{1}{N_i}\mathbf{z}_i\mathbf{z}_i^T$, and N_i is the number of elements in the set ω_i and $\mathbf{z}_i = [z_{i1}, z_{i2}, \ldots, z_{in}]^T$, $i = 1, \ldots, M$, $N_1 + \cdots + N_M = N$. Now we take the vector ϕ_j such that

$$\mathbf{R}\phi_j = \lambda_j \phi_j,$$

for λ_j which are the real nonnegative eigenvalues of the matrix R. After sorting the eigenvalues decreasingly ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq \lambda_n$) we can take the normed eigenvectors $\epsilon_1, \ldots, \epsilon_m$ and define the matrix ϕ of the form:

$$\phi = \begin{bmatrix} \epsilon_1^T \\ \vdots \\ \epsilon_m^T \end{bmatrix},$$

where m is the dimension of the new space Z^*. Transformed output data $\mathbf{x}$ we get from the equation:

$$\mathbf{x} = \phi\mathbf{y}.$$

For the output data we can apply a method of classification which bases on the Bayes classifier and the maximum likelihood estimator. Let us consider M multivariate normally distributed classes with parameters $\mathbf{m}_i \in \mathbb{R}^m$ and $\mathbf{c}_i \in \mathbb{R}^m_m$, $i = 1, \dots, M$. According to the density function let

$$p(\mathbf{x}, \omega_i) = \frac{1}{(2\pi)^{m/2}|\mathbf{c}_i|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m}_i)^T \mathbf{c}_i^{-1}(\mathbf{x} - \mathbf{m}_i)\right), \ \mathbf{x} \in \mathbb{R}^m. \quad (1)$$

The likelihood function is given by Bayes classifier:

$$d_i(\mathbf{x}) = p(\mathbf{x}, \omega_i)p_r(\omega_i), \quad i = 1, \dots, M.$$

If we take the equation (1) in the one above and the estimates of $\mathbf{m}_i$, $\mathbf{c}_i$, $p_r(\omega_i)$ instead of their unknown values, and additionally after taking the logarithmic function of it and missing a constant we will have

$$\tilde{d}_i(\mathbf{x}) = \ln\tilde{p_r}(\omega_i) - \frac{1}{2}\ln|\tilde{\mathbf{c}}_i| - \frac{1}{2}(\mathbf{x} - \tilde{\mathbf{m}}_i)^T \tilde{\mathbf{c}}_i{}^{-1}(\mathbf{x} - \tilde{\mathbf{m}}_i),$$

where

$$\tilde{\mathbf{m}}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} \mathbf{x}_j, \quad (2)$$

$$\tilde{\mathbf{c}}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} \mathbf{x}_j\mathbf{x}_j^T - \tilde{\mathbf{m}}_i\tilde{\mathbf{m}}_i{}^T \quad (3)$$

are the maximum likelihood estimates. A classifying function which separates the classes ω_i and ω_j (a separating surface) is given by

$$\tilde{d}_i(\mathbf{x}) - \tilde{d}_j(\mathbf{x}) = 0. \quad (4)$$

In the case of 2-dimensional space ($m = 2$ and $\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$) the separating surface (4) can be written in the form

$$ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f = 0. \quad (5)$$

As it will be shown in our article, it is worth considering also other estimators instead of (2) and (3): robust estimators based on the Huber's functions.

3. An estimator based on the Huber's functions

In the article [11] authors consider a problem of a robust estimation for the family of distributions $N_m(A\mu, \Sigma)$, where $A \in \mathbb{R}^m_s$ is a known matrix with s independent columns, a vector $\mu = (\mu_1, \dots, \mu_s)^T \in \mathbb{R}^s$ and a positive definite matrix $\Sigma \in \mathbb{R}^m_m$ are unknown parameters. In the case when $s = m$ and

$A = \mathbb{I}_m$ is the identity matrix we get the family of distributions $N_m(\mu, \Sigma)$ with unknown parameters μ and Σ. The authors present a method of a robust estimation that is based on statistical functionals.

Let $\mathcal{G}$ denote a set of cumulative distribution functions $F\colon \mathbb{R}^m \to [0,1]$ and let Θ be a parameter space.

Definition 3.1 A function defined for $F \in \mathcal{G}$ and taking values in the parameter space Θ is called a statistical functional.

For a sample $\mathbf{X}_1, \ldots, \mathbf{X}_N$ an estimator of a parameter $\theta \in \Theta$ can be defined as

$$\hat{\theta} = T(\hat{F_N}), \tag{6}$$

where T is a statistical functional and $\hat{F_N}\colon \mathbb{R}^m \to [0,1]$ is the empirical cumulative distribution function of form

$$\hat{F_N}(t_1, \ldots, t_m) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{(-\infty, t_1] \times \cdots \times (-\infty, t_m]}(\mathbf{X}_i), \quad (t_1, \ldots, t_m)^T \in \mathbb{R}^m.$$

Example 3.2 Let $T^*\colon \mathcal{G} \to \Theta$ be the function defined by

$$T^*(G) = \operatorname*{argmin}_{\theta \in \Theta} \int \Phi(\mathbf{x}|\theta) \mathrm{d}G(\mathbf{x}), \tag{7}$$

where $\Phi\colon \mathbb{R}^m \times \Theta \to \mathbb{R}$ is a given function and the right side of the equation (7) denotes the parameter θ for which the function $\int \Phi(\mathbf{x}|\theta)\mathrm{d}G(\mathbf{x})$ attains the minimum. The function T^* is a statistical functional.

In the article [11] authors present the unknown matrix Σ as $\Sigma = \sum_{i=1}^{k} \alpha_i W_i$, where the matrices $W_1, \ldots, W_k$, $k = \frac{m(m+1)}{2}$, are a basis of the space of real square and symmetric matrices. The parameters $\alpha_1 \in \mathbb{R}, \ldots, \alpha_k \in \mathbb{R}$ are unknown. Denote by $\mu = (\mu_1, \ldots, \mu_m)^T$ and $\alpha = (\alpha_1, \ldots, \alpha_k)^T$. The parameter space Θ can be written in the form

$$\Theta = \left\{ \theta = \begin{pmatrix} \mu \\ \alpha \end{pmatrix} \in \mathbb{R}^{m+k} : \mu \in \mathbb{R}^m \wedge \alpha \in \mathbb{R}^k \wedge \sum_{i=1}^{k} \alpha_i W_i > 0 \right\}.$$

The authors use the statistical functional of form (7) with the objective function Φ of form

$$\Phi(\mathbf{x}|\theta) = \ln |\sum_{i=1}^{k} \alpha_i W_i|^{\frac{1}{2}} + \varphi \left(\frac{1}{c_\varphi^2} (\mathbf{x} - \mu)^T \left(\sum_{i=1}^{k} \alpha_i W_i \right)^{-1} (\mathbf{x} - \mu) \right), \tag{8}$$

where $c_\varphi > 0$ is a constant and $\varphi\colon [0,+\infty) \to \mathbb{R}$ is a function. For $\varphi(u) = \frac{1}{2}u$ and $c_\varphi = 1$ the function Φ is the function that corresponds to the maximum likelihood estimator.

Generally, the function φ and the constant c_φ should be properly chosen. It can be done according to the following theorem and the conditions (B1)-(B4) which are given below.

Theorem 3.1 Let ξ be a random vector having the multivariate standard normal distribution $N_m(0, \mathbb{I}_m)$ and assume that a function $\varphi\colon [0,+\infty) \to \mathbb{R}$ satisfies the following conditions.

(B1) The function φ has positive derivative on $(0,+\infty)$.
(B2) The function $u\varphi'(u^2)$ has nonnegative derivative on $[0,+\infty)$ and there exists $u_0 > 0$ such that $2u_0^2\varphi'(u_0^2) > m$.

Then there exists a unique $c_{\varphi,m} > 0$ for which the function

$$c \longmapsto m\ln(c) + \mathbb{E}\left[\varphi\left(\frac{\xi^T\xi}{c^2}\right)\right],\ c > 0,$$

attains the global minimum.

A proof of the theorem can be found in the article of S. Zontek [14].

Assume that the function $\varphi\colon [0,+\infty) \to \mathbb{R}$ satisfies the conditions (B1), (B2) and also

(B3) The function φ'' is continuous.
(B4) The functions $u\varphi'(u^2)$ and $u^2\varphi''(u^2)$ are bounded.

The authors consider the statistical functional T^* given by (7) with the objective function Φ of form (8) for the constant $c_\varphi = c_{\varphi,m}$. The estimator (6) is then robust.

4. An attempt of choosing the most effective estimator for the motors problem

In the article [1] the authors considered data connected with motors of type SZXb6514 B made by Zakład Silników Elektrycznych Małej Mocy "Silma" in Sosnowiec (Low Power Electric Motors Company "Silma"). They considered 11 usable motors and 23 motors which are not usable. The not usable motors were grouped with respect to the following types of defects:

- B - rubbing,
- C - loudness - load operation,
- E - high current,
- F - increased vibration level,

- G - no rivet in the sheet package.

The motors were represented by 9-dimensional vectors of the features $x_{1A}(x, y, z)$, $x_{2v}(x, y, z)$, $x_{3a}(x, y, z)$ (Table 1). Using the KLT the authors reduced each of the three 3-dimensional features to one point and got 3-dimensional vectors of the features $x'_1(A)$, $x'_2(v)$, $x'_3(a)$ (Table 2 on the left). The 3-dimensional space was reduced to a 2-dimensional space (x''_1, x''_2) by the KLT (Table 2 on the right).

Table 1: The motors input data.

Amplitude x_{1A}			Velocity x_{2v}			Acceleration x_{3a}			
x	y	z	x	y	z	x	y	z	
2.2	1.9	3	1.5	1.35	2.2	3	3	5.5	A
1.4	1	3.1	1	0.6	2.5	3.5	2.5	6.5	A
1.8	1.15	4	1.25	0.9	3	3	2.3	7	A
2.3	1.9	3.5	1.7	1.6	2.7	4	4	8	A
1.2	2	1.8	2	1.45	4	2	2.1	4.7	A
1.3	1	3	2	1.7	8	1.8	1.2	7	A
2.3	1.8	2.1	4	3	3.8	3.6	1.6	2.2	A
2.5	4.5	2.6	1.7	3.5	2	4.1	2.6	6	A
4	2.6	5.6	1.9	2	4	2.6	2.2	4.1	A
2.6	2.5	3.6	2.5	2.1	5.1	3.1	2.6	5.1	A
1.5	1.7	3.2	2.5	1.3	1.45	2.6	3	6	A
1.9	1.2	14.5	1.3	1.1	9.5	13.5	10.5	32	B
4	3.9	7.3	5.5	4.5	9	7.2	7	4	B
2.3	2.2	7	6.5	4.5	5.3	10	10.5	3.5	B
1.5	1	5	1.25	0.8	3	23	14	30	C
1.9	1.6	3.3	1.6	1.35	2.5	13.5	9	14	C
2	1.55	1.4	1.4	1.15	3	19	24	30	C
2.3	1.75	1.1	1.7	1.4	2.5	20.5	10	14.5	C
1.9	1.2	5.5	1.3	1.1	9.5	13.5	10.5	22	C
4.5	4	20.5	4	6	13	15	15.5	30	E
17	14	19	1.1	9	40	15.5	18.5	35	E
9	6.4	30	7	4.5	20.5	7	7	26	E
4	3.5	18	7	7	35	5.7	5.5	35	E
8	7	22.1	6	10	27	14.5	14.5	29	E
30	32	17	2.5	2	11.5	10	15	35	E
14	24.5	14.5	3.7	6	10	9	16	29	E
19	17.5	60	15	14	40	18	14	35	F
27.5	26.1	22.1	22	17.4	18	11.8	9	14.9	F
7.5	5.5	20.5	5.5	3	14	6.2	5.5	15	G
7	5	18	5.5	3	13	7.5	5.5	14	G
7.1	5.4	19	5.8	3.2	13.5	7	6	15.1	G
7.4	5.5	19.5	5.6	3.1	14	6.8	5.8	15.1	G
7.3	5.6	20	5.7	3.2	14	6.2	5.6	15	G
9.8	9	15.5	2.8	4.9	12.5	7	12.5	12.5	G

The following classification problems with two classes were considered:

- USABLE motors (A) - UNUSABLE motors (BCEFG),
- B - CEFG,
- C - BEFG,
- E - BCFG,
- F - BCEG,
- G - BCEF.

The probability $\tilde{p_r}(\omega)$ of the occurrence of a usable object according to the manufacturer's information is set at 0.98. In the case of testing the type of defect, it was assumed that the occurrence of the above failure classes is equally probable with the probability $\tilde{p_r}(\omega)$ of 0.5.

The Figure 2 presents the image of the 2-dimensional output data which was obtained by KLT.

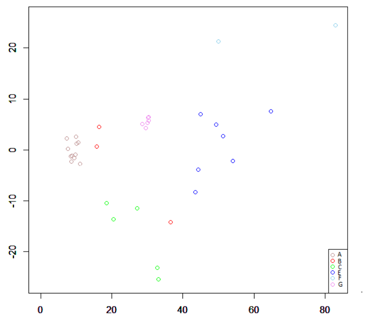

Figure 2: The 2-dimensional output motors data.

The classifying functions (4) for the following estimators are compared:

- the maximum likelihood estimator MLE,
- the robust estimator $\hat{\theta}_t$ given by (6) and the statistical functional of form (7), where the objective function Φ is defined by (8) and the function $\varphi\colon [0, +\infty) \to \mathbb{R}$ given by

$$\varphi(x) = \tilde{\phi}_t(\sqrt{x}), \quad x \geq 0,$$

where the function $\tilde{\phi}_t$ is defined by its derivative

$$\tilde{\phi}_t'(x) = \begin{cases} x, & |x| \leq t, \\ -x - 4t - \frac{2t^2}{x}, & -2t < x < -t, \\ -x + 4t - \frac{2t^2}{x}, & t < x < 2t, \\ \frac{2t^2}{x}, & |x| \geq 2t. \end{cases}$$

for a given t and the corresponding constant c given in Table 3.

Table 2: The motors data after the first (on the left) and the second KLT (on the right).

$x_1'(A)$	$x_2'(v)$	$x_3'(a)$	
4.144	2.869	6.936	A
3.498	2.749	7.768	A
4.452	3.369	7.895	A
4.586	3.457	9.796	A
2.795	4.677	5.513	A
3.375	8.331	6.969	A
3.429	5.623	3.921	A
5.047	3.392	7.644	A
7.293	5.14	5.306	A
5.048	6.02	6.492	A
3.915	3.412	7.182	A
12.88	9.263	36	B
9.192	11.22	9.004	B
7.5	8.244	11.12	B
5.055	3.339	39.64	C
4.127	3.171	20.61	C
2.636	3.493	41.92	C
2.612	3.219	24.3	C
5.712	9.263	27.77	C
27.92	11.96	33.83	E
40.02	11.72	38.79	E
20.39	13.57	18.07	E
19.97	14.74	36.93	E
28.41	42.14	42.43	E
30.5	22.02	27.01	E
17.55	35.73	33.29	E
24.02	29.13	35.51	F
63.41	44.95	41.69	F
40.52	28.6	20.65	G
21.9	15.25	17.05	G
19.48	14.36	16.76	G
20.49	14.97	17.66	G
21.06	15.32	17.49	G
21.46	15.38	17.09	G

x_1''	x_2''	
8.49	-1.243	A
8.658	-2.252	A
9.522	-1.567	A
11.05	-2.676	A
7.635	0.1889	A
10.62	1.429	A
7.392	2.246	A
9.826	-0.9381	A
10	2.578	A
10.12	1.279	A
8.822	-1.17	A
36.56	-14.14	B
16.5	4.5	B
15.8	0.7246	B
32.75	-23.19	C
18.49	-10.49	C
33.13	-25.42	C
20.51	-13.64	C
27.14	-11.41	C
43.44	-8.309	E
64.7	7.559	E
44.93	7.05	E
49.35	4.921	E
51.22	2.697	E
54.07	-2.199	E
44.36	-3.852	E
82.83	24.44	F
50.04	21.29	F
30.37	6.432	G
28.58	5.111	G
30	5.29	G
30.32	5.822	G
30.25	6.311	G
29.57	4.328	G

Table 3: The constants t and c.

t	1.1	1.3	1.5	1.7	2	2.5
c	0.373	0.669	0.768	0.810	0.836	0.844

Using the classifying function given by (5) we get different separating surfaces for the each considered case. The separating surfaces are presented in the following six figures (Figures 3 - 8). It is easy to see that the obtained shapes of the separating surfaces often differ radically. This is not surprising because each of these images graphically describes a different type of damage.

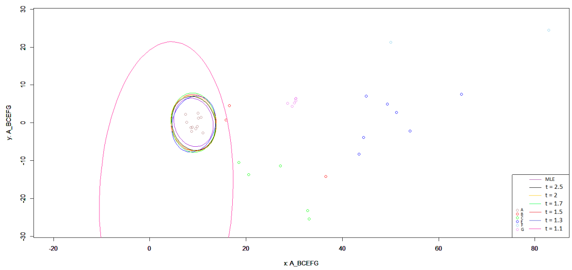

Figure 3: The separating surfaces for the case A - BCEFG.

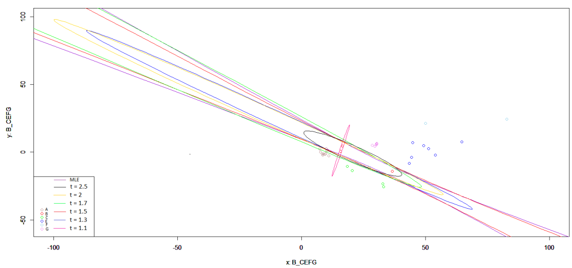

Figure 4: The separating surfaces for the case B - CEFG.

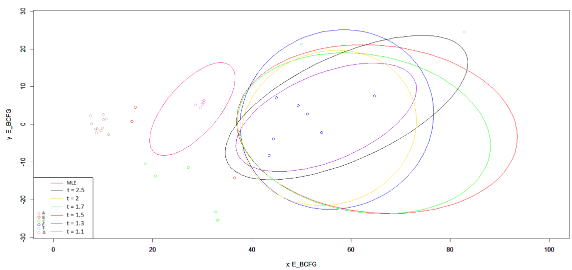

Figure 5: The separating surfaces for the case E - BCFG.

The above figures describe the separating surfaces corresponding to six training sets: A-BCEFG, B-CEFG, C-BEFG, E-BCFG, F-BCEG and G-BCEF. Let us assume that a new motor with unknown classification is subjected to the recognition process. By measuring the amplitude of displacement, velocity and acceleration of vibrations in three axes, as in the case of input data (the training set) we form a 9-dimensional vector, which then through the KLT we transpose into 2-dimensional space according to eigenvectors obtained for the corresponding training set.

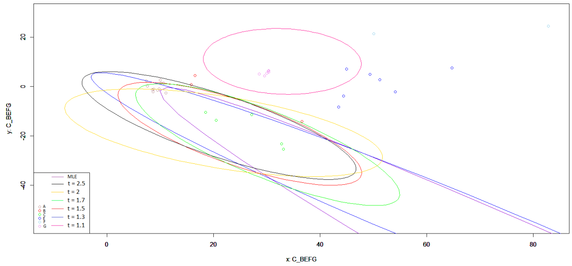

Figure 6: The separating surfaces for the case C - BEFG.

Table 4: The percentages of wrongly classified elements for all cases.

	$t = 1.1$	$t = 1.3$	$t = 1.5$	$t = 1.7$	$t = 2$	$t = 2.5$	MLE
A - BCEFG	0%	0%	0%	0%	0%	0%	0%
B - CEFG	13%	13%	17.4%	17.4%	17.4%	8.7%	13%
C - BEFG	47.8%	8.7%	4.3%	4.3%	8.7%	4.3%	4.3%
E - BCFG	39.1%	4.3%	13%	8.7%	13%	17.4%	17.4%
F - BCEG	0%	4.8%	4.8%	0%	0%	4.8%	0%
G - BCEF	4.3%	0%	4.3%	0%	0%	0%	4.3%
The average	17.4%	5.1%	7.3%	5.1%	6.5%	5.9%	6.5%

The decision recognition system reviews these partitioning surfaces and determines the state class based on belonging to the area to which the unknown object was assigned. This allows to specify the two-state classification USABLE - UNUSABLE, and then in the case of including the object in the class UNUSABLE - to define the type of defect of the tested motor. The process of assigning a class label to a new engine involves the possibility of making a mistake. The Table 4 shows the percentages of wrongly classified elements to compare the decision functions corresponding to the considered estimators (the Leave-One-Out method, see T. Hastie at al. [7]). The Leave-One-Out method is a type of cross-validation and consists of dividing the N-element set into N one-element subsets. Each subset in turn becomes the test set, and the rest together form the training set. As we can see the use of a Huber's function may reduce the probability of making a mistake in a classification.

5. Concluding remarks and an open problem The cited considerations were made based on incomplete training sets. The obtained results are, however, encouraging to undertake further research, because as the method has shown it allows to determine areas typical for a given object state class. It could be an interesting open problem to find an optimal Huber's function (a

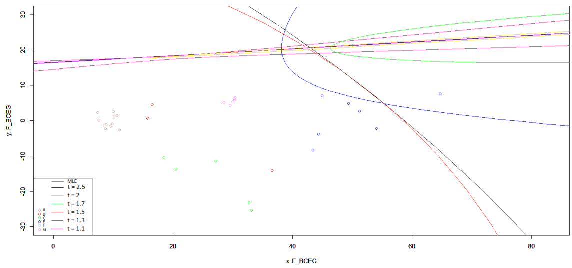

Figure 7: The separating surfaces for the case F - BCEG.

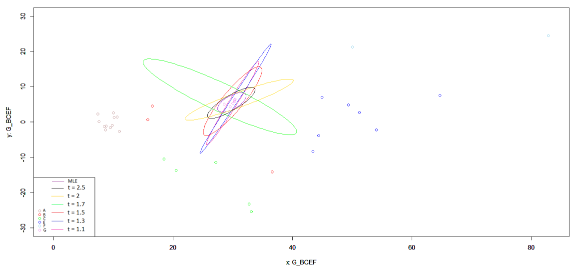

Figure 8: The separating surfaces for the case G - BCEF.

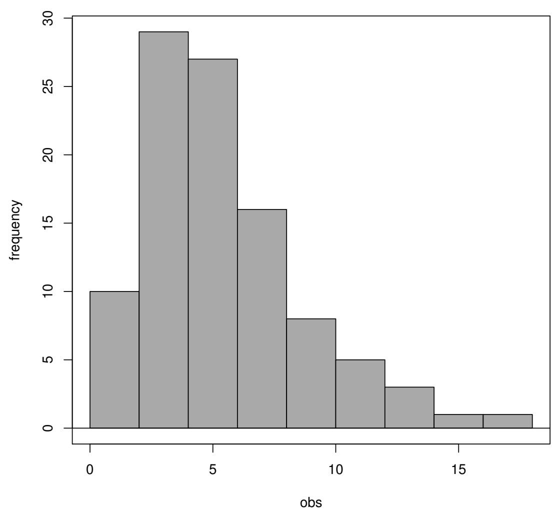

Figure 9: Przykład.

tunning constant t) for given data which minimizes the probability of making a mistake.

Acknowledgments

The Authors are indebted to the Reviewers for their valuable remarks and editorial help.

6. References

[1] J. Adamczewski and H. Gacki. Próba zastosowania metod rozpoznawania obrazów w eksperymencie diagnostycznym. *PAN Zagadnienia eksploatacji maszyn*, 29(97):161–172, 1994. No cited.

[2] T. Bednarski and S. Zontek. Robust estimation of parameters in a mixed unbalanced model. *Ann. Statist.*, 24(4):1493–1510, 1996. No cited.

[3] B. Desgraupes. conics: Plot conics, "r" package version 0.3, 2013. No cited.

[4] H. Gacki and A. Kulawik. Robust estimation and its application to a classification problem. *Mathematica Applicanda*, 47(2):81–94, 2019. No cited.

[5] R. G. Gonzalez and J. Z. Tou. *Pattern recognition principles.* Addison-Wesly Publ. Comp, 1974. No cited.

[6] V. Goulet, C. Dutang, M. Maechler, D. Firth, M. Shapira, and M. Stadelmann. expm: Matrix exponential, log, 'etc', "r" package version 0.999-4, 2019. No cited.

[7] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning, Data Mining, Inference and Prediction.* Springer, 2017. No cited.

[8] P. J. Huber. *Robust Statistics.* Wiley, New York, 1981. No cited.

[9] F. Jager. Feature extraction and shape representation of ambulatory electrocardiogram using the karhunen-loeve transform. *Electrotechnical Review*, 69(2):83–89, 2002. No cited.

[10] A. Kulawik and S. Zontek. Robust estimation in the multivariate normal model with variance components. *Statistics*, 49(4):766–780, 2015. No cited.

[11] A. Kulawik and S. Zontek. Robust estimation in the multivariate normal model. *Discussiones Mathematicae Probability and Statistics*, 36(1-2): 53–66, 2016. No cited.

[12] M. Nieniewski. Probabilistic methods in pattern recognition and computer vision. *Lecture Notes and Computer Vision and Artificial Intelligence, Ossolineum*, 1990. No cited.

[13] R. Zmyślony and S. Zontek. Robust m-estimator of parameters in variance component model. *Discussiones Mathematicae Probability and Statistics*, 22(1-2):61–71, 2002. No cited.

[14] S. Zontek. Multivariate robust estimation in linear model for spatially located sensors and random input. *Discussiones Mathematicae Algebra and Stochastic Methods*, 18:195–206, 1998. No cited.

Funkcje Hubera i ich zastosowanie w pewnym problemie klasyfikacji

Henryk Gacki, Agnieszka Kulawik

Streszczenie W artykule rozważany jest problem klasyfikacji w przypadku dwóch klas o wielowymiarowym rozkładzie normalnym. Problem ten jest rozwiązywany na podstawie przykładu empirycznego (dane dotyczące silników) z wykorzystaniem transformacji Karhunena-Loevego oraz funkcji klasyfikujących bazujących na wybranych estymatorach nieznanych parametrów wielowymiarowego rozkładu normalnego. Rozważany jest zarówno klasyczny estymator - estymator największej wiarogodności, jak również estymator odporny, który opiera się o funkcje Hubera. Uzyskane klasyfikatory są porównywane za pomocą sprawdzianu krzyżowego - metoda Leave-One-Out.

Klasyfikacja tematyczna AMS (2010): 62C12; 62P30.

Słowa kluczowe: funkcje Hubera, wielowymiarowy rozkład normalny, estymator odporny, klasyfikator, transformacja Karhunena-Loevego.

Henryk Gacki[a] was awarded a Master of Science diploma in Mathematics in 1976, the University of Silesia, PhD degree in mathematics from the University of Silesia in 1984 and DSC (habilitation) degree in 2008 in mathematics from the University of Silesia. In 2016 Professorship at the University of Silesia in Katowice. In the years 2016 – 2019 Chair of the Section of Mathematical Methods in Economy and Finances, the Institute of Mathematics, the University of Silesia. In the years 2012 – 2015 Coordinator of Mathematics as a code of Modernity European Programme under whose framework a target field of study was being realized at the University of Silesia. Since 2015 Intermaths Polish–Italian Master's Double Degree Programme Coordinator at the University of Silesia. For the year 2014 Statistical editor in "Chowanna". A member of the Polish Mathematical Association. Scientific interests: Theory of semigroups of Markov operators and their applications, testing stability of solutions of different types of the Boltzmann equation on the space of signed measures, generalizations and applications of the Kantorovich–Rubinstein maximum principle constituting an element of the solution in the transport of mass problem, probability theory and statistics. Distinctions and awards: 2016 Golden Badge of Merit for the University of Silesia, 2018 Golden Honorary Badge for Merits for the Silesian Voivodship, 2018 Selected Athlete of the 50th anniversary of AZS the University of Silesia.

[a]References to his research papers can be found in MathSciNet under ID:223263 and the EMS, FIZ Karlsruhe, and the Heidelberg Acad. Sci. bibliography database (zbMath) under ai:Gacki.Henryk.

Agnieszka Kulawik received the MSc degree in Mathematics from the University of Silesia in Katowice (2007) and PhD degree in Mathematics from the University of Zielona Góra (2011). She is interested in the following research topics: mathematical statistics and application of statistics. References to her research papers are found in MathSciNet under ID:1115942 and the European Mathematical Society, FIZ Karlsruhe, and the Heidelberg Academy of Sciences bibliography database known as zbMath under ai:Kulawik.Agnieszka.

Henryk Gacki
University of Silesia in Katowice, Faculty of Science and Technology
Bankowa 14, 40-007 Katowice, Poland
E-mail: henryk.gacki@us.edu.pl

Agnieszka Kulawik
University of Silesia in Katowice, Faculty of Science and Technology
Bankowa 14, 40-007 Katowice, Poland
E-mail: agnieszka.kulawik@us.edu.pl

Communicated by: Marek Kałuszka

(Received: 19th of July 2020; revised: 9th of October 2020)