Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

# Bayes Parameter Identification with Polynomial Asymmetrical Loss Function

Piotr Kulczycki*, Aleksander Mazgaj**

* Systems Research Institute, Polish Academy of Sciences, Poland
*(e-mail: kulczycki@ibspan.waw.pl , http://www.ibspan.waw.pl/kulczycki)*
** Cracow University of Technology, Department of Automatic Control, Poland

**Abstract:** The parameter identification for problems where losses arising from overestimation and underestimation are different and can be described by an asymmetrical and polynomial function, is investigated here. The Bayes decision rule allowing to minimize potential losses is used. Calculation algorithms are based on the nonparametric methodology of statistical kernel estimators, which frees the method from distribution type. Three basic cases are considered in detail: a linear, a quadratic, and finally a general concept for a higher order polynomial – here the cube-case is described in detail as an example. For each of them the final result constitutes a numerical procedure enabling to effectively calculate the optimal value of a parameter in question.

## 1. INTRODUCTION

Besides classic or trivial cases, the creation of an ideal model for an object under automatic control is neither possible, nor even required, as it would be far too complicated for effective use (Morrison, 1991; Nusse and Yorke, 1997; Soderstrom and Stoica, 1994). Consequently, absolutely precise determination of the values of parameters contained within is impossible, not only from a metrological point of view, but also due to the fact that such a value does not even exist, while a considered parameter represents an entire range of phenomena impossible to describe in a form of a single number. As identification is in practice always subject to a higher goal (usually conditioned by the control algorithm), then more suitable results can be obtained thanks to the consideration, in the estimation of the parameters' values, of the losses implied through errors encountered here. Often such losses can be described by the function assuming the following asymmetrical and polynomial form:

$$l(\hat{x}, x) = \begin{cases} (-1)^k \, a\,(\hat{x}-x)^k & \text{for} \quad \hat{x}-x \le 0 \\ b\,(\hat{x}-x)^k & \text{for} \quad \hat{x}-x \ge 0 \end{cases} , \tag{1}$$

with $k \in \mathbb{N} \setminus \{0\}$, where the coefficients $a$ and $b$ are positive, and may differ, while $x$ and $\hat{x}$ denote the parameter under consideration and its estimator, respectively.

Consider therefore the typical situation when one has $m$ values of the investigated parameter $x_1$, $x_2, \dots, x_m$, obtained by measuring or directly with the aid of auxiliary quantities. In this paper, the uncertainty of the examined parameter is considered with a probabilistic approach. The nonparametric methodology of statistical kernel estimators will be applied to identify the distribution of probability measure, which makes the result independent of arbitrary assumptions concerning the type of this distribution. An algorithm based on the Bayes decision rule is investigated, which allows to obtain minimal expectation value for potential losses. The proposed procedure is universal and can be applied in a wide range of tasks, not only in the field of engineering. Furthermore the worked out method can be used for other uncertainty approaches apart from that of probability, e.g. fuzzy logic.

Three basic cases will be investigated in the following: linear (Section 3.1), quadratic (Section 3.2), and higher order polynomial (Section 3.3) – here the cube-case will be described in detail. In every case the final result will be an algorithm for the calculation of values for an optimal estimator, ensuring that its practical implementation does not demand of the user detailed knowledge of the theoretical aspects, or laborious research and calculations.

## 2. MATHEMATICAL PRELIMINARIES

### *2.1. Bayes Decision Rule*

The main aim of decision theory is the selection of a concrete decision based only on a representation of measure characterizing the imprecision of states of nature. The one-dimensional case will be considered in the following. Let there be given the nonempty set of states of nature $Z = \mathbb{R}$, and the nonempty set of possible decisions $D \subset \mathbb{R}$. Assume that the imprecision of states of nature is of probability type and its distribution is described by the density $f : \mathbb{R} \to [0, \infty)$. Let there be given also the loss function $l : D \times Z \to \mathbb{R} \cup \{\pm\infty\}$, while its values $l(d, z)$ can be interpreted as losses occurring in a hypothetical case, when the state of nature is $z$ and the decision $d$ is taken. If for every $d \in D$ the integral $\int_{\mathbb{R}} l(d, z)\, f(z)\, \mathrm{d}z$ exists, then the Bayes loss function $l_B : D \to \mathbb{R} \cup \{\pm\infty\}$ can be defined as

$$l_B(d) = \int_{\mathbb{R}} l(d,z)\, f(z)\, \mathrm{d}z \quad . \tag{2}$$

Every element $d_B \in D$ such that $l_B(d_B) = \min_{d \in D} l_B(d)$ is called a Bayes decision, and the above procedure – a Bayes decision rule. The Bayes decision minimizes the mean value of losses following the decision $d$. Further details are found in (Berger, 1980).

### *2.2. Statistical Kernel Estimators*

Let the one-dimensional random variable *X*, with a distribution having the density *f*, be given. Its kernel estimator $\hat{f} : \mathbb{R} \to [0,\infty)$ is calculated on the basis of the *m*-element random sample $x_1$, $x_2, ..., x_m$ acquired experimentally from the variable *X*, and is defined in its basic form by the formula

$$\hat{f}(x) = \frac{1}{mh} \sum_{i=1}^{m} K\left(\frac{x - x_i}{h}\right) \quad , \tag{3}$$

where the function $K : \mathbb{R} \to [0,\infty)$ measurable, symmetrical relative to zero, with a weak global maximum at this point, and fulfilling the condition $\int_{\mathbb{R}} K(x)\, \mathrm{d}x = 1$, is called a kernel, whereas the positive coefficient $h$ is known as a smoothing parameter. It should be stated that the kernel estimators allow the identification of density of practically any distribution, without an arbitrary assumption concerning its type.

Fixing values introduced in definition (3), i.e. choosing the form of the kernel *K* and calculating the smoothing parameter *h* value, is mostly carried out using the mean square criterion.

Thus, from the statistical point of view, the form of the kernel seems not to have essential meaning, thanks to which it becomes possible for the choice of the function *K* to be arbitrary, taking into account above all required properties of the estimator obtained, e.g. class of regularity, positive values, or other qualities important in the case of a particular problem, especially the convenience of calculations.

As opposed to the form of the kernel, the value of the smoothing parameter has significant influence on the quality of the estimator obtained, but fortunately many convenient algorithms have been developed. For the one-dimensional case considered here, the direct plug-in method is strongly recommended – for details see (Kulczycki, 2005, Section 3.1.5; Wand and Jones, 1995, Section 3.6.1).

Practical tasks call for the application of various useful procedures, which generally improve the quality of estimation, as well as other – facultative – suiting a model to fit the reality under research. For the needs of the problem considered hereinafter, modification of the smoothing parameter (Kulczycki, 2005, Section 3.1.6; Silverman, 1986, Section 5.3.1) is strongly recommended from the first type of procedures, while from the latter, the boundaries of a support of continuous variable $X$ (Kulczycki, 2005, Section 3.1.8; Silverman, 1986, Section 2.10) may be also applied.

Detailed descriptions of the statistical kernel estimators methodology are found in (Kulczycki, 2005; Silverman, 1986; Wand and Jones, 1995) as well as – with exemplary applications in systems research and industry – in (Kulczycki, 2007, 2008).

## 3. ALGORITHM

### *3.1. Linear Case*

As an example illustrating the investigations presented in this section, an optimal control problem will be considered. Such systems have shown themselves in practice to be very sensitive to the inaccuracy of modeling, which was – in fact – the main limit of their applications. The control performance index which exists here, however, can also refer to quality of identification allowing the creation of an optimal procedure for estimation of model parameter values, thereby notably lowering this sensitivity.

Thus, consider the following dynamic system:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \dfrac{1}{m} \end{bmatrix} u(t) \quad , \tag{4}$$

where the positive parameter $m$ represents a mass submitted to a force according to Newton's second law of dynamics. Then $x_1$, $x_2$ and $u$ denotes position and velocity of the mass, and the force regarded here as a control, respectively. Such a system constitutes a basis for the majority of research in the field of robotics, leading in consequence to much more complex models, specifically suited to the particular problem under investigation. Consider the time-optimal control task, the basic form of which consists of bringing the system's state to the origin, in minimal and finite time, assuming the control values are bounded. For details see the classic textbook (Athans and Falb, 1966, Chapter 7). Fundamental meaning for phenomena existing in the control system lies in proper identification of value of the parameter $m$. The control is defined in relation to the value of the estimator $\hat{m}$, actually different from the value of the parameter $m$ in the object. Detailed analysis is found in (Kulczycki and Wisniewski, 2002).

In the purely hypothetical case of $\hat{m} = m$, i.e. when the value of the estimator of this parameter is equal to its true value, the process is regular in character. The system's state reaches the origin in minimal and finite time. However, in the event of overestimation (i.e. $\hat{m} < m$), over-regulations occur in the system – its state oscillates around the origin and reaches it in a finite time, albeit larger than the minimal. Next, in the case of underestimation (i.e. when $\hat{m} > m$), the system's state moves along a so-called sliding trajectory and finally reaches the origin in a finite time, again larger than the minimal. Figure 1 shows the graph of the performance index for values of the estimator $\hat{m}$. One can note that an increase in this index is roughly proportional to the estimation error $|\hat{m}-m|$, although with different coefficients for positive and negative errors. The resulting losses can so be described in the form of an asymmetrical linear loss function, i.e. given by formula (1) with $k=1$.

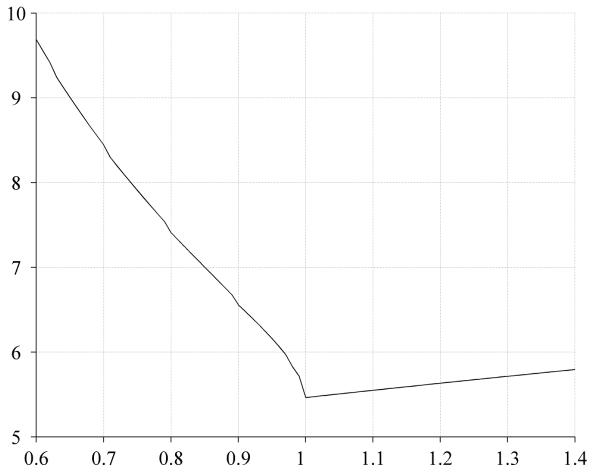


Fig. 1. Performance index $J$ for values of the estimator $\hat{m}$, with $m=1$

The parameter under investigation, whose value is to be estimated, will be denoted by $x$. In order to adhere to the principles of decision theory presented in Section 2.1, it will be treated here as the value of a random variable. According to point estimation methodology, it is assumed that the metrologically achieved measurements of the above parameter, i.e. $x_1$, $x_2, \ldots, x_m$, are the sum of its "true" (although unknown) value and random disturbances of various origin. The goal of this research is the calculation of the estimator of this parameter (hereinafter denoted by $\hat{x}$), which would approximate the "true" value – the best from the point of view of a practical problem investigated. In order to solve this task, the Bayes decision rule will be used, ensuring a minimum of expectation value of losses. According to the conditions formulated above, the loss function is assumed in asymmetrical linear form:

$$l(\hat{x},x)=\begin{cases}-a(\hat{x}-x) & \text{for} \quad \hat{x}-x\le 0\\ b(\hat{x}-x) & \text{for} \quad \hat{x}-x\ge 0\end{cases} , \tag{5}$$

while the coefficients $a$ and $b$ are positive and not necessarily equal to each other. Thus, the Bayes loss function (2) is given by the formula

$$l_B(\hat{x})=b\int_{-\infty}^{\hat{x}}(\hat{x}-x)f(x)\,\mathrm{d}x-a\int_{\hat{x}}^{\infty}(\hat{x}-x)f(x)\,\mathrm{d}x \quad , \tag{6}$$

where $f:\mathbb{R}\to[0,\infty)$ denotes the density of distribution of a random variable representing the uncertainty of states of nature, i.e. the parameter in question. It is readily shown that the function $l_B$ fulfils its minimum for the value being a solution of the following equation with the argument $\hat{x}$:

$$\int_{-\infty}^{\hat{x}}f(x)\,\mathrm{d}x-\frac{a}{a+b}=0 \quad . \tag{7}$$

Since $0<a/(a+b)<1$, a solution for the above equation exists, and if the function $f$ has connected support, e.g. it is positive, this solution is unique. Moreover, thanks to equality

$$\frac{a}{a+b}=\frac{\frac{a}{b}}{\frac{a}{b}+1} \quad , \tag{8}$$

it is not necessary to identify the parameters $a$ and $b$ separately, rather only their ratio.

The identification of the density $f$ present in condition (7) will be carried out using statistical kernel estimators, presented in Section 2.2. Then one should choose a continuous kernel of positive values and also so that the function $I:\mathbb{R}\to\mathbb{R}$ such that $I(x)=\int_{-\infty}^{x}K(y)\,\mathrm{d}y$ can be expressed by relatively simple analytical formula. In consequence, this results in a similar property regarding the function $U_i:\mathbb{R}\to\mathbb{R}$ for any fixed $i=1,2,\ldots,m$ defined as

$$U_i(x)=\frac{1}{h}\int_{-\infty}^{x}K\left(\frac{y-x_i}{h}\right)\mathrm{d}y \quad . \tag{9}$$

Then criterion (7) can be expressed equivalently in a form of

$$\frac{h}{m}\sum_{i=1}^{m}U_i(\hat{x})-\frac{a}{(a+b)}=0 \quad . \tag{10}$$

If the left side of the above formula is denoted by $L(\hat{x})$, its derivative is simply

$$L'(\hat{x})=\hat{f}(\hat{x}) \quad , \tag{11}$$

where $\hat{f}$ was given by definition (3). In this situation, the solution of criterion (7) can be calculated numerically on the basis of Newton's algorithm (Kinkaid and Cheney, 2002) as the limit of the sequence $\{\hat{x}_j\}_{j=0}^{\infty}$ defined by

$$\hat{x}_0=\frac{1}{m}\sum_{i=1}^{m}x_i \tag{12}$$

$$\hat{x}_{j+1} = \hat{x}_j - \frac{L(\hat{x}_j)}{L'(\hat{x}_j)} \quad \text{for} \quad j = 0,1,\ldots \quad , \tag{13}$$

with the functions $L$ and $L'$ being given by formulas (10)-(11), whereas a stop criterion takes on the form

$$|\hat{x}_j - \hat{x}_{j-1}| \le 0{,}01\hat{\sigma} \quad , \tag{14}$$

where $\hat{\sigma}$ denotes the estimator of the standard deviation obtained from the sample $x_1$, $x_2, \ldots, x_m$.

In the linear case worked out above, the Cauchy kernel

$$K(x) = \frac{2}{\pi} \frac{1}{(1+x^2)^2} \tag{15}$$

is proposed. Then

$$U_i(x) = \frac{1}{\pi} \operatorname{arctg}\left(\frac{x-x_i}{h}\right) + \frac{\frac{x-x_i}{h}}{\pi\left[1+\left(\frac{x-x_i}{h}\right)^2\right]} + \frac{1}{2} \quad . \tag{16}$$

*3.2. Quadratic Case*

As an example, consider the problem concerning the classical task of optimal control for a quadratic performance index (Athans and Falb, 1966, Section 9.5) with infinite end time and unit matrix/parameter for the integrand function of the performance index. The object is the dynamic system

$$\dot{x}(t) = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \lambda \end{bmatrix} u(t) \quad , \tag{17}$$

where $\lambda \in \mathbb{R} \setminus \{0\}$. Moreover, let $\hat{\lambda} \in \mathbb{R} \setminus \{0\}$ represent an estimator of the parameter $\lambda$. An optimal feedback controller is defined on the basis of the value $\hat{\lambda}$, not necessarily equal to the value of the parameter $\lambda$ existing in the object. The values of the performance index obtained for a particular $\hat{\lambda}$, are shown in Fig. 2. One can see that the resulting graph can be described with great precision by a quadratic function with different coefficients for positive and negative errors, which in fact proves that over- and underestimation of the parameter $\lambda$ have other results on the performance index value.

To use an analogous methodology to that of the linear case considered in the previous section, the loss function is assumed in quadratic and asymmetrical form defined as

$$l(\hat{x}, x) = \begin{cases} a(\hat{x}-x)^2 & \text{for} \quad \hat{x}-x \le 0 \\ b(\hat{x}-x)^2 & \text{for} \quad \hat{x}-x \ge 0 \end{cases} \quad , \tag{18}$$

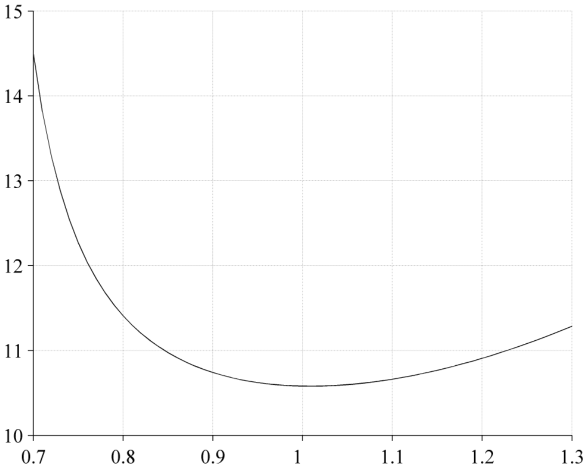


Fig. 2. Performance index $J$ for values of the estimator $\hat{\lambda}$, with $\lambda = 1$

while the coefficients $a$ and $b$ are positive and not necessarily equal to each other. Thus, the Bayes loss function (2) is given by the formula

$$l_B(\hat{x}) = a\int_{\hat{x}}^{\infty} (\hat{x}-x)^2 f(x)\,\mathrm{d}x + b\int_{-\infty}^{\hat{x}} (\hat{x}-x)^2 f(x)\,\mathrm{d}x \quad . \tag{19}$$

One can show that the function $l_B$ fulfils its minimum for the value $\hat{x}$ being a solution of the equation

$$(a-b)\int_{-\infty}^{\hat{x}} (\hat{x}-x) f(x)\,\mathrm{d}x - a\int_{-\infty}^{\infty} (\hat{x}-x) f(x)\,\mathrm{d}x = 0 \quad . \tag{20}$$

This solution exists and is unique. Like in the linear case, dividing the above equation by $b$, note that it is necessary to identify only the ratio of the parameters $a$ and $b$.

Solution of equation (20) for a general case is not an easy task. However, if estimation of the density $f$ is reached using statistical kernel estimators, then – thanks to a proper choice of the kernel form – one can design an effective numerical algorithm to this end. Let, therefore, a continuous kernel of positive values, fulfilling the condition

$$\int_{-\infty}^{\infty} xK(x)\,\mathrm{d}x < \infty \tag{21}$$

be given. Besides the functions $U_i$ introduced in Section 3.1, let for any fixed $i = 1, 2, \ldots, m$ the functions $V_i : \mathbb{R} \to \mathbb{R}$ be defined as

$$V_i(x) = \frac{1}{h} \int_{-\infty}^{x} yK\left(\frac{y-x_i}{h}\right) \mathrm{d}y \quad . \tag{22}$$

The kernel $K$ should be chosen so the function $J:\mathbb{R}\to\mathbb{R}$ such that $J(x)=\int_{-\infty}^{x} y\,K(y)\,\mathrm{d}y$ be expressed by a convenient analytical formula.

If an expected value is estimated by the arithmetical mean value of a sample, then criterion (20) can be described equivalently as

$$\sum_{i=1}^{m}[(a-b)(\hat{x}U_i(\hat{x})-V_i(\hat{x}))+ax_i]-a\hat{x}m=0 \quad . \tag{23}$$

If the left side of the above formula is denoted by $L(\hat{x})$, then – using the equality $V_i'(\hat{x})=\hat{x}U_i'(\hat{x})$ directly resulting from dependencies (9) and (22) – one can express the value of its derivative as

$$L'(\hat{x})=\sum_{i=1}^{m}[(a-b)U_i(\hat{x})]-am \quad . \tag{24}$$

In this situation, the solution of criterion (20) can be calculated numerically on the basis of Newton's algorithm (12)-(14). In the quadratic case also Cauchy kernel (15) is proposed; then formula (16) remains true and additionally:

$$V_i(x)=x_i\left(\frac{1}{\pi}\operatorname{arctg}\left(\frac{x-x_i}{h}\right)+\frac{\frac{x-x_i}{h}}{\pi\left[1+\left(\frac{x-x_i}{h}\right)^2\right]}+\frac{1}{2}\right)-\frac{h}{\pi\left[1+\left(\frac{x-x_i}{h}\right)^2\right]} \quad . \tag{25}$$

*3.2. Polynomial Case*

In this section, detailed investigations presented earlier will be supplemented with the polynomial case, that is where the loss function is an asymmetrical monomial of the order $k\geq 2$ and is therefore given by the following formula:

$$l(\hat{x},x)=\begin{cases}(-1)^k a(\hat{x}-x)^k & \text{for } \hat{x}-x\leq 0\\ b(\hat{x}-x)^k & \text{for } \hat{x}-x\geq 0\end{cases} \quad , \tag{26}$$

while the coefficients $a$ and $b$ are positive, and may differ. Criterion for the optimal estimator $\hat{x}$ is given here in the form

$$(-1)^k ak\int_{\hat{x}}^{\infty}(\hat{x}-x)^{k-1}f(x)\,\mathrm{d}x+bk\int_{-\infty}^{\hat{x}}(\hat{x}-x)^{k-1}f(x)\,\mathrm{d}x=0 \quad . \tag{27}$$

The solution of the above equation exists and is unique.

When the statistical kernel estimators are used with respect to the density $f$, it is possible again to create an efficient numerical algorithm enabling equation (27) to be solved. Let the kernel $K$ be continuous, of positive values and fulfilling the following condition:

$$\int_{-\infty}^{\infty}x^{k-1}K(x)\,\mathrm{d}x<\infty \quad . \tag{28}$$

For clarity of presentation, the case $k=3$ is presented below. Thus, equation (27), after simple transformations, takes on the equivalent form

$$(a+b)\left(\hat{x}^2\int_{-\infty}^{\hat{x}}f(x)\,\mathrm{d}x-2\hat{x}\int_{-\infty}^{\hat{x}}xf(x)\,\mathrm{d}x+\int_{-\infty}^{\hat{x}}x^2f(x)\,\mathrm{d}x\right)-a\left(\hat{x}^2-2\hat{x}\int_{-\infty}^{\infty}xf(x)\,\mathrm{d}x+\int_{-\infty}^{\infty}x^2f(x)\,\mathrm{d}x\right)=0 \quad . \tag{29}$$

Now, with any fixed $i=1,2,\ldots,m$, let the functions $U_i$ and $V_i$ defined by dependencies (9) and (22) be given, and furthermore $W_i:\mathbb{R}\to\mathbb{R}$ be introduced as

$$W_i(x)=\frac{1}{h}\int_{-\infty}^{x}y^2K\left(\frac{y-x_i}{h}\right)\mathrm{d}y \quad . \tag{30}$$

Making use of the above notations, condition (28) can be expressed in the following form

$$\sum_{i=1}^{m}\left[(a+b)\left(x^2U_i(x)-2xV_i(x)+W_i(x)\right)+2ax_ix-\lim_{x\to\infty}W_i(x)\right]-amx^2=0 \quad . \tag{31}$$

If the left-hand side of the above formula is denoted as $L(x)$, then – also taking into account the equalities $V_i'(x)=xU_i'(x)$ and $W_i'(x)=xV_i'(x)$ resulting from dependencies (9), (22) and (30) – the derivative of the function $L$ is

$$L'(x)=\sum_{i=1}^{m}\left[2(a+b)\left(xU_i(x)-V_i(x)\right)+2ax_i\right]-2amx \quad . \tag{32}$$

Finally, the desired estimator can be calculated numerically through Newton's algorithm (12)-(14), while the functions $L$ and $L'$ are given by formulas (31)-(32).

The Cauchy kernel (15) must by modified here to the form

$$K(x) = \frac{8}{3\pi} \frac{1}{(1+x^2)^3} \quad . \tag{33}$$

An increase of the power in the denominator has been implied with the necessity of ensuring the fulfillment of condition (28). Here:

$$U_i(x) = \frac{3\left(\frac{x-x_i}{h}\right)^3 + 5\frac{x-x_i}{h}}{3\pi\left(1+\left(\frac{x-x_i}{h}\right)^2\right)^2} + \frac{1}{\pi}\operatorname{arctg}\left(\frac{x-x_i}{h}\right) + \frac{1}{2} \tag{34}$$

$$V_i(x) = -\frac{2}{3\pi}\frac{h}{\left[1+\left(\frac{x-x_i}{h}\right)^2\right]^2} + x_i\left(\frac{3\left(\frac{x-x_i}{h}\right)^3 + 5\frac{x-x_i}{h}}{3\pi\left[1+\left(\frac{x-x_i}{h}\right)^2\right]^2} + \frac{1}{\pi}\operatorname{arctg}\left(\frac{x-x_i}{h}\right) + \frac{1}{2}\right) \tag{35}$$

$$W_i(x) = -\frac{4hx_i}{3\pi\left[1+\left(\frac{x-x_i}{h}\right)^2\right]^2} + x_i^2\left(\frac{3\left(\frac{x-x_i}{h}\right)^3 + 5\frac{x-x_i}{h}}{3\pi\left[1+\left(\frac{x-x_i}{h}\right)^2\right]^2} + \frac{1}{\pi}\operatorname{arctg}\left(\frac{x-x_i}{h}\right) + \frac{1}{2}\right) + h^2\left(\frac{\left(\frac{x-x_i}{h}\right)^3 - \frac{x-x_i}{h}}{3\pi\left[1+\left(\frac{x-x_i}{h}\right)^2\right]^2} + \frac{1}{3\pi}\operatorname{arctg}\left(\frac{x-x_i}{h}\right) + \frac{1}{6}\right) \quad . \tag{36}$$

The above investigations can be analogously transposed to a higher order of asymmetrical polynomial loss function (1), although on account of their extreme nature, they seem to be useful mainly for untypical applicational tasks.

## 4. NUMERICAL SIMULATION RESULTS

The operation of the algorithm designed here has been checked in detail using a numerical simulation. In the case $a = b$, the results were close to medium value, however, when $a \neq b$, the algorithm provided possibilities that cannot be achieved using classical methods, by appropriately shifting the value of the estimator in the direction associated with smaller losses, where intensity of this process was stimulated by the parameter $k$ depending on the nature of the system under research. Many different distributions were examined including e.g. multimodal with asymmetrical modes. In each case, as the size of a random sample $m$ increases, the mean estimation error and its standard deviation tend to zero. From an applicational point of view, these fundamental properties are demanded of estimators used in practice. This, above all, states that as the sample size increases, the estimators' values achieved tend to the desired value, and their dispersion decreases. This allows for the obtaining of any required precision, although the proper sample size must be guaranteed. In practice this implies a necessity for compromise between these two quantities. A satisfactory degree of precision was obtained when the size of the sample was between 10 and 200, i.e. for $m \in [10, 200]$ ; in particular, the large values became necessary when the difference between parameters *a* and *b* increased.

Generally, the benefits arising from application of the method proposed in this paper are greater the more complex the control system is, and over- and under-estimation of a model's parameters have a more differing influence on performance index, i.e. when asymmetry of the loss function is more distinct.

## REFERENCES

Athans, M. and Falb, P.L. (1994). *Optimal Control*. McGraw-Hill, New York.

Berger, J.O. (1980). *Statistical Decision Theory*. Springer-Verlag, New York.

Kinkaid, D. and Cheney, W. (2002). *Numerical Analysis*. Brooks/Cole, Pacific Grove.

Kulczycki, P. (2005). *Estymatory jadrowe w analizie systemowej*. WNT, Warsaw.

Kulczycki, P. (2007). Estymatory jadrowe w badaniach systemowych. In: *Techniki informacyjne w badaniach systemowych* (Kulczycki, P., Hryniewicz, O., Kacprzyk, J. (Ed)), 79-105, WNT, Warsaw.

Kulczycki, P. (2008). Kernel Estimators in Industrial Applications. In: *Soft Computing Applications in Industry* (Prasad, B. (Ed)), 69-91, Springer-Verlag, Berlin.

Kulczycki, P. and Wisniewski, R. (2002). Fuzzy Controller for a System with Uncertain Load. *Fuzzy Sets and Systems*, **131**, 185-195.

Morrison, F. (1991). *The Art of Modeling Dynamic Systems*, Multiscience Press, New York.

Nusse, H.E. and Yorke, J.A. (1997). *Dynamics: Numerical Exploration*, Springer-Verlag, New York.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, New York.

Soderstrom, T.S. and Stoica, P. (1994) *System Identyfication*, Prentice Hall, Englewood Cliffs.

Wand, M.P. and Jones, M.C. (1995). *Kernel Smoothing*. Chapman and Hall, London.