

Computational Aspects of a Method of Stochastic Approximation

Konstantin Runovski, Igor Rystsov and Hans-Jürgen Schmeisser

Abstract. A method of stochastic approximation is studied in the framework of the general convergence theory for families of linear polynomial operators of interpolation type. The description of the corresponding computational procedure, in particular, its input parameters, is given. Some optimization problems and aspects of implementation of the algorithm by means of *Maple* are discussed. It is shown that the algorithm can be applied not only to problems of "pure approximation" in the spaces L_p with $0 < p \leq +\infty$, but also to problems of signal processing, especially, if one is interested in strong oscillating data or data containing an essential stochastic item.

Keywords. Fast Fourier transform, random numbers, families of linear polynomial operators, approximation algorithms

Mathematics Subject Classification (2000). 41A05, 42A10, 42A61, 42B08, 65D05, 65D15

1. Introduction

In this paper we discuss a method and an algorithm of stochastic approximation, which are based on approximations by families of linear polynomial operators given on the L_p -space of functions defined on the d -dimensional torus \mathbb{T}^d with $0 < p \leq +\infty$ by setting

$$\mathcal{L}_{n;\lambda}^\varphi(f, x) = (2n + 1)^{-d} \sum_{|k| \leq n} f(t_n^k + \lambda) W_n(x - t_n^k - \lambda) \quad (f \in L_p, n \in \mathbb{N}_0). \quad (1.1)$$

Here $\lambda \in \mathbb{R}^d$ is a parameter,

$$W_0(h) \equiv 1, \quad W_n(h) = \sum_{|k| \leq n} \varphi\left(\frac{k}{n}\right) e^{ikh} \quad (n \in \mathbb{N}), \quad (1.2)$$

K. Runovski: Analytical Center of Investment, 01001 Kiev, Ukraine;
murom@lycos.com

I. Rystsov: National Technical University, 01001 Kiev, Ukraine;
mozart@voliacable.com

H.-J. Schmeisser: Mathematical Institute, Friedrich-Schiller-University, 07737 Jena, Germany;
mhj@minet.uni-jena.de

$\varphi(\cdot)$ is a real valued centrally symmetric continuous function with support contained in the unit ball of \mathbb{R}^d satisfying $\varphi(0) = 1$, and

$$t_n^k = \frac{2\pi k}{2n+1} \quad (k \in \mathbb{Z}^d). \quad (1.3)$$

The families (1.1) were systematically studied in [1] and [4] – [6]. The main property our method is based on is that the averaged approximation error in L_p , $0 < p \leq +\infty$, with respect to the parameter λ can be estimated up to a constant by the best approximation of order $\asymp n$, if the generator φ of the kernel satisfies the following additional conditions: $\varphi(\xi) = 1$ at the neighborhood of 0, and its Fourier transform belongs to $L_{\tilde{p}}(\mathbb{R}^d)$, where $\tilde{p} = \min(1, p)$. One of the possible versions of this result we need in this paper will be given in Section 2 (Theorem 1). We also show that the same quality of approximation can be achieved, if the parameter λ is randomly chosen. This result enables us to elaborate the procedure we have called *algorithm of stochastic approximation (SA-algorithm)*, which reduces the problem of trigonometric approximation for all admissible parameters $0 < p \leq +\infty$ to the problem of interpolation with randomly shifted nodes (Theorem 2).

Formula (1.1) is very close in a certain sense to the formulas determining the classical methods of trigonometric approximation, namely, the interpolation means and the Fourier means. More precisely, the interpolation means are given by (1.1) with $\lambda = 0$, and the Fourier means are obtained, if $\lambda = 0$ again, and the summation over the discrete set of points t_n^k is replaced by the integration over the continuous parameter. For the complete comparison of approximation properties of all three methods we refer to [6]. Here we notice that the families of linear polynomial operators of interpolation type can be used for approximation in all L_p -spaces with $0 < p \leq +\infty$ in contrast to the interpolation means and the Fourier means which are relevant only for $p = +\infty$ and $1 \leq p \leq +\infty$, respectively.

From the computational point of view the approximation algorithm based on the interpolation means (*I-algorithm*) is much more economical in comparison with the procedure based on the Fourier means (*F-algorithm*). Indeed, the calculation of $\mathcal{L}_{n,0}(f, x)$ is reduced to the calculation of the discrete Fourier coefficients (the Fourier-Lagrange coefficients) of f , that is, to the Fast Fourier Transform (*FFT*). Because of this reason the I-algorithm consists of $O((n \log n)^d)$ operations, where n is the order of the approximation polynomial in the sense of (1.2), in contrast to the F-algorithm, where the number of operations is essentially larger, mainly, because of necessity to compute the Fourier coefficients, that is, integrals from functions oscillating at least up to n times. Indeed, by means of applying standard cubatures, such a computation is reduced to the FFT as well. However, as the simplest test for $\cos nx$ shows, the number of nodes N , that is, the length of the sequence to be transformed with the help of

the FFT, one needs to achieve any acceptable accuracy, should be much greater than n .

The SA-algorithm we discuss in this paper enables us to keep the simplicity of the I-algorithm, which is applied to continuous functions only, for all admissible parameters $0 < p \leq +\infty$. It should be noticed in this respect that we have not found in the literature any other computational procedure for approximation in the L_p -spaces with $0 < p < 1$. As it was already mentioned above such a combination of "nice properties" is achieved by random translations of interpolation nodes. These random shifts can be controlled by an additional parameter $0 < \varepsilon < 1$ which we call *probability error*. In the limiting case of approximation by interpolation means it is equal to 0.

The paper is organized as follows. In Section 2 we formulate and prove the basic theorem on stochastic approximations. The discussion on the input parameters and their relations are given in Section 3. Section 4 contains the computational procedure. Some optimization problems are posed and discussed in Section 5. Some aspects of implementation of the SA-algorithm on the *Maple*-platform are described in Section 6. Section 7 contains the results applying the SA-algorithm to test functions and a discussion of possible applications and further developments. In particular, the problem of detecting a main analytical component of a process whose development partially depends on the influence of some uncontrolled random effects, as well as the ways of applying the SA-algorithm to its solution are described.

2. Theorem on stochastic approximation

Our starting point is a result on the quality of the approximation by families of type (1.1)–(1.2) which can be easily obtained following the ideas presented in [6] (see, in particular, (2.8) and Theorem 7.7). As usual,

$$E_\sigma(f)_p = \inf_{T_\sigma} \|f - T_\sigma\|_p \quad (f \in L_p, \sigma \geq 0),$$

where the infimum is taken over all real valued trigonometric polynomials with spectrum contained in the ball of radius σ centered at the origin, is *the best approximation to f of order σ in the space L_p* . By $\mathcal{G}(q, \rho)$, where $0 < q \leq 1$, $0 < \rho < 1$, we denote the class of real valued centrally symmetric continuous functions φ satisfying

$$\varphi(\xi) = \begin{cases} 1, & |\xi| \leq \rho \\ 0, & |\xi| > 1, \end{cases}$$

whose Fourier transform $\widehat{\varphi}$ belongs to $L_q(\mathbb{R}^d)$. We use the symbol $L_{\tilde{p}}$ for the L_p -space of functions of two variables x and λ , which is equipped with the norm

$$\|\cdot\|_{\tilde{p}} = \begin{cases} (2\pi)^{-\frac{d}{p}} \|\cdot\|_{p;x} \|_{p;\lambda}, & 0 < p < +\infty \\ \|\cdot\|_{p;x} \|_{p;\lambda}, & p = +\infty. \end{cases}$$

where $\|\cdot\|_{p;z}$ denotes the (quasi-)norm with respect to the variable z .

Theorem 1. *Let $0 < p \leq +\infty$, $0 < \rho < 1$ and $\varphi \in \mathcal{G}(\tilde{p}, \rho)$, where $\tilde{p} = \min(1, p)$. Then*

$$\|f - \mathcal{L}_{n;\lambda}^\varphi(f)\|_{\tilde{p}} \leq c(d, p, \varphi) E_{\rho n}(f)_p \quad (f \in L_p, n \in \mathbb{N}_0), \tag{2.1}$$

where

$$c(d, p, \varphi) = \begin{cases} 1 + (2\pi)^{-d} 3^{d(1-\frac{1}{p})} \|\widehat{\varphi}\|_{L_1(\mathbb{R}^d)}, & 1 \leq p \leq +\infty \\ \left(1 + (2\pi)^{-d} 3^{d(1-p)} \|\widehat{\varphi}\|_{L_p(\mathbb{R}^d)}^p\right)^{\frac{1}{p}}, & 0 < p < 1 \end{cases} \tag{2.2}$$

Theorem 1 shows that the approximation by families of linear polynomial operators of interpolation type gives almost the best order of approximation in average on the parameter λ for $0 < p < +\infty$. Combining this result with the classical Chebyshev inequality we will show that the same outcome can be obtained with an a priori defined probability error, if the parameter λ is randomly chosen, or, in other words, the interpolation nodes are randomly shifted without loss of uniformity of the grid. The one-dimensional version of the theorem below without estimates for the constant can be found in [1]. Henceforth, the probability of an event A is denoted with $P\{A\}$. We use also the symbol \mathcal{E} for the mathematical expectation.

Theorem 2 (Basic theorem on stochastic approximation). *Let $\gamma > 1$, $0 < p < +\infty$, $0 < \rho < 1$, $\varphi \in \mathcal{G}(\tilde{p}, \rho)$, where $\tilde{p} = \min(1, p)$. Let also $m \in \mathbb{N}$ and η_j , $j = 1, \dots, m$, be independent random vectors uniformly distributed on the unit cube $[0, 1]^d$. Then for $f \in L_p$ and $n \in \mathbb{N}_0$*

$$P \left\{ \min_{j=1, \dots, m} \|f - \mathcal{L}_{n;\theta_j}^\varphi(f)\|_p \leq \gamma c(d, p, \varphi) E_{\rho n}(f)_p \right\} \geq 1 - \gamma^{-pm}, \tag{2.3}$$

where $\theta_j = \tau \eta_j$, $j = 1, \dots, m$, and $\tau = \frac{2\pi}{(2n+1)}$.

Proof. Taking into account that the random vectors θ_j are independent and

that they have the same distribution law we get ($c \equiv c(d, p, \varphi)$)

$$\begin{aligned}
 & P\left\{\min_{j=1,\dots,m} \|f - \mathcal{L}_{n;\theta_j}^\varphi(f)\|_p \leq \gamma c E_{\rho n}(f)_p\right\} \\
 &= 1 - P\left\{\bigcap_{j=1}^m \left\{\|f - \mathcal{L}_{n;\theta_j}^\varphi(f)\|_p > \gamma c E_{\rho n}(f)_p\right\}\right\} \\
 &= 1 - \prod_{j=1}^m P\left\{\|f - \mathcal{L}_{n;\theta_j}^\varphi(f)\|_p > \gamma c E_{\rho n}(f)_p\right\} \quad (2.4) \\
 &= 1 - \left(P\left\{\|f - \mathcal{L}_{n;\theta_1}^\varphi(f)\|_p > \gamma c E_{\rho n}(f)_p\right\}\right)^m \\
 &= 1 - \left(P\left\{\|f - \mathcal{L}_{n;\theta_1}^\varphi(f)\|_p^p > (\gamma c E_{\rho n}(f)_p)^p\right\}\right)^m.
 \end{aligned}$$

By Chebyshev's inequality we have

$$P\left\{\|f - \mathcal{L}_{n;\theta_1}^\varphi(f)\|_p^p > (\gamma c E_{\rho n}(f)_p)^p\right\} \leq (\gamma c E_{\rho n}(f)_p)^{-p} \mathcal{E}(\|f - \mathcal{L}_{n;\theta_1}^\varphi(f)\|_p^p). \quad (2.5)$$

Since

$$dF_{\theta_1}(\lambda) = \tau^{-d} d\lambda \quad (\lambda \in [0, \tau]^d),$$

where F_{θ_1} is the distribution function for θ_1 and, as it follows immediately from (1.1), the function $\|f - \mathcal{L}_{n;\lambda}^\varphi(f)\|_p$ is τ -periodic on each λ_i , $i = 1, \dots, d$, we get

$$\begin{aligned}
 \mathcal{E}(\|f - \mathcal{L}_{n;\theta_1}^\varphi(f)\|_p^p) &= \tau^{-d} \int_{[0,\tau]^d} \|f - \mathcal{L}_{n;\lambda}^\varphi(f)\|_p^p d\lambda \\
 &= (2\pi)^{-d} \int_{\mathbb{T}^d} \|f - \mathcal{L}_{n;\lambda}^\varphi(f)\|_p^p d\lambda \quad (2.6) \\
 &= \|f - \mathcal{L}_{n;\lambda}^\varphi(f)\|_p^p.
 \end{aligned}$$

Combining (2.4)–(2.6) we obtain

$$P\left\{\min_{j=1,\dots,m} \|f - \mathcal{L}_{n;\theta_j}^\varphi(f)\|_p \leq \gamma c E_{\rho n}(f)_p\right\} \geq 1 - \left((\gamma c E_{\rho n}(f)_p)^{-p} \|f - \mathcal{L}_{n;\lambda}^\varphi(f)\|_p^p\right)^m.$$

Now (2.3) follows immediately from (2.1). \square

3. Input parameters

Now we discuss the input parameters of the SA-algorithm. They can be split into three groups: given, operating and dependent parameters. The parameters of the first group are determined by the problem to be solved. We present them in the following table.

	Name	Range
d	dimension	\mathbb{N}
p	metric	$(0, +\infty)$
f	function to be approximated	$L_p(\mathbb{T}^d)$
σ	order of best approximation	$[0, +\infty)$
ε	probability error	$(0, 1)$

Table 1. Given parameters

In contrast to the first group the operating parameters can be chosen according to our decision. Some optimization problems caused by this "freedom" will be discussed later. The parameters of the second group are gathered in the Table 2.

	Name	Range
ρ	relative order of best approximation to the order of kernel	$(0, 1)$
φ	generator of the kernel	$\mathcal{G}(\tilde{p}, \rho)$
γ	constant of exceeding "theoretical level"	$(1, +\infty)$

Table 2. Operating parameters

The depending parameters are presented in the Table 3.

	Name	Range
n	order of the kernel	\mathbb{N}_0
m	number of random shifts	\mathbb{N}

Table 3. Depending parameters

In view of (1.2) n is a non-negative integer. In the multivariate case n means the radius of the ball on the lattice \mathbb{Z}^d , and the requirement to be integer does not look very natural. However, the number $2n + 1$ has to be an integer even in this case, in order to guarantee the periodicity of the set of interpolation nodes (see (1.3)). Thus, the parameter n , if it is not 0, has to be an integer or a number of type $s - \frac{1}{2}$, where s is an integer. Sometimes, as we shall see in

Section 6, the second possibility turns out to be more preferable. In contrast to n the parameter σ can be any real non-negative number.

Since the sequence of best approximations is decreasing, the order of the kernel, that is, the order of the approximating polynomial, should be chosen in such a way that the parameter of the best approximation ρn will be as close as possible to the prescribed σ . By this remark we get

$$n = \min \{k : k\rho \geq \sigma\}, \tag{3.1}$$

where k is s or $s - \frac{1}{2}$, $s \in \mathbb{N}$.

In order to be able to apply the FFT-procedure, we have to require that the number $2n + 1$, that is, the number of nodes of the grid, is a power of 2. In view of (3.1) this implies that

$$n = 2^{r-1} - \frac{1}{2}, \quad r = \min \left\{ q \in \mathbb{N} : q \geq \log_2 \left(\frac{2\sigma}{\rho} + 1 \right) \right\} \tag{3.2}$$

In order to attain the required probability precision the right-hand side of (2.3) has to be greater than $1 - \varepsilon$, that is,

$$\gamma^{-pm} < \varepsilon. \tag{3.3}$$

Taking into account that the parameter m , which influences the number of operations, should be chosen as small as possible, we get from (3.3) the explicit formula for the optimal number of random shifts of the grid

$$m = \left\lceil \frac{\log \frac{1}{\varepsilon}}{p \log \gamma} \right\rceil + 1. \tag{3.4}$$

4. Computational procedure

In this section we describe the procedure of stochastic approximation in more detail.

In *the first step* the random independent uniformly distributed on $[0, 1]^d$ vectors η_j , $j = 1, \dots, m$, where m is calculated by means of formula (3.4), should be generated. For this purpose standard programs and tools of applied stochastics can be used. Then we determine $\theta_j = \tau\eta_j$, $j = 1, \dots, m$.

In *the second step* we apply the Fast Fourier Transform (see, for instance, [2] or [8] for more details) to the sets of the values of a given function on the shifted uniform grids

$$\begin{aligned} F_j &= \left\{ f(t_n^\nu + \theta_j) : \nu \in \mathbb{Z}^d, 0 \leq \nu_i \leq 2n, i = 1, \dots, d \right\} \\ &\downarrow \text{FFT} \\ \widehat{F}_j &= \left\{ c_k^{(n)}(f_j) : k \in \mathbb{Z}^d, 0 \leq k_i \leq 2n, i = 1, \dots, d \right\}, \end{aligned} \tag{4.1}$$

$j = 1, \dots, m$, where

$$c_k^{(n)}(f_j) = (2n + 1)^{-d} \sum_{\nu_1=0}^{2n} \dots \sum_{\nu_d=0}^{2n} f_j(t_n^\nu) e^{-ik(t_n^\nu + \theta_j)} \quad (k \in \mathbb{Z}^d), \quad (4.2)$$

are the Fourier-Lagrange coefficients of the function $f_j = f(\cdot + \theta_j)$.*) By (4.2) and (1.1) we obtain

$$\mathcal{L}_{n;\theta_j}^\varphi(f; x) = \sum_{|k| \leq n} \varphi\left(\frac{k}{n}\right) c_k^{(n)}(f_j) e^{ikx} \quad (j = 1, \dots, m). \quad (4.3)$$

In the third step we calculate the errors

$$\|f - \mathcal{L}_{n;\theta_j}^\varphi(f)\|_p \quad (j = 1, \dots, m), \quad (4.4)$$

and choose the minimal one. If it corresponds to the index j_0 , then by Theorem 2 the polynomial $\mathcal{L}_{n;\theta_{j_0}}^\varphi(f)$ approximates a given function f up to the approximation error $\gamma c(d, p, \varphi) E_{\rho n}(f)_p$ with probability not less than $1 - \varepsilon$.

Finally, we notice that we apply the formula (1.1) m times, finding the minimum of m numbers requires m operations and the number of operations one needs to produce m random values is linear dependent on m . Thus, the complete SA-algorithm has the same property.

5. Optimization problems

In this Section we give some remarks concerning the optimal choice of the parameters and we pose the corresponding optimization problems.

The quality of the computational procedure above is mainly determined by the following characteristics:

- nearness to the best possible theoretical approximation order (γ, ρ) ;
- reliability of the outcome (ε) ;
- number of operations (m) ,

where in brackets the main responsible parameters are given. Formula (3.4) shows that the improvement of any of these characteristics can be achieved on account of the other ones. Which parameter should be given the preference mainly depends on the specific characters of the problem to be solved.

The parameters ρ and φ do not influence the properties of the SA-algorithm itself. On the other hand, they are responsible for the quality of estimate (2.1). We are not able to give the complete theoretical solution of the problem of their

*) Clearly, the function $c_k^{(n)}$ given on \mathbb{Z}^d has a period $2n + 1$ with respect to each k_i .

optimal choice. However, we discuss some general principles which enable us to find quasi-optimal values by computer experiments.

The parameter ρ is present in both the constant $c(d, p, \varphi)$ and in the value $E_{\rho n}(f)_p$. In view of (2.2) the behavior of $c(d, p, \varphi)$ is determined by the norm of the Fourier transform of the generator of the kernel. Let us consider the case $1 \leq p \leq +\infty$. Using that $\varphi(0) = 1$ we obtain a trivial estimate from below

$$\|\widehat{\varphi}\|_{L_1(\mathbb{R}^d)} \geq 1. \tag{5.1}$$

Taking into account that in the one-dimensional case the value 1 is attained, for instance, for the generator of the classical Fejer means, that is, for $\rho = 0$, and that the norm of the generator of the Fourier partial sums, for which $\rho = 1$, is infinity, we can conjecture that the smaller ρ the better the constant $c(d, p, \varphi)$. On the other hand, the smaller ρ the greater the value $E_{\rho n}(f)_p$. These two trends which contradict each other determine the choice of ρ .

Before we comment on these observations in more detail, we notice that, if ρ is somehow chosen, it remains the possibility to vary the function φ in the ring $\rho < |\xi| < 1$. In view of (2.2) we can formulate the associated optimization problem as follows. *Given $0 < p \leq 1$ and $0 < \rho < 1$. It is required to find a function $\varphi_* \in \mathcal{G}(p, \rho)$ giving the solution of the problem:*

$$\|\widehat{\varphi}\|_p \longrightarrow \min \quad (\varphi \in \mathcal{G}(p, \rho)). \tag{5.2}$$

Continuing the discussion on the choice of ρ and φ we demonstrate by the example below that the optimal value of ρ depends on the smoothness properties of the given function f . Let $d = 1$, $p = +\infty$ and let f be smooth of order $\alpha > 0$. This condition can be understood in various meanings. Here we suppose that f has continuous derivatives up to the order $[\alpha]$, so that $f^{([\alpha])}$ satisfies a Hölder condition with the parameter $\{\alpha\} = \alpha - [\alpha]$. By the version of classical Jackson estimate for smooth functions (see, for instance, [7, pp. 287–292]) one has

$$E_n(f) \leq \frac{CM(f)}{(\rho n)^\alpha} \quad (n \in \mathbb{N}), \tag{5.3}$$

where C depends only on α and

$$M(f) = \sup_{h>0} (h^{-\{\alpha\}} \|\Delta_h f^{([\alpha])}(x)\|).$$

We consider the function

$$\varphi_\rho(\xi) = \begin{cases} 1, & |\xi| \leq \rho \\ \frac{\rho - |\xi|}{1 - \rho} + 1, & \rho < |\xi| \leq 1 \\ 0, & |\xi| > 1. \end{cases} \tag{5.4}$$

By direct calculation we have

$$\widehat{\varphi}_\rho(x) = \frac{4 \sin \frac{1-\rho}{2}x \sin \frac{1+\rho}{2}x}{(1-\rho)x^2}. \quad (5.5)$$

Using (5.5) it is easy to check that

$$\|\widehat{\varphi}_\rho\|_1 \asymp (1-\rho)^{-1}. \quad (5.6)$$

By (2.2), (5.3) and (5.6) we obtain for the right-hand side of the inequality (2.1)

$$c(1, +\infty, \varphi_\rho) E_{\rho n}(f) \leq C' M(f) (\rho(1-\rho))^{-\alpha} n^{-\alpha},$$

where C' is an absolute constant. The function $(\rho(1-\rho))^{-\alpha}$ has a minimum on the interval $(0, 1)$, which is attained at the point

$$\rho_* = \frac{\alpha}{\alpha + 1}. \quad (5.7)$$

In particular, for the class of methods generated by functions of type (5.4) the classical Vallée-Poussin mean ($\rho = \frac{1}{2}$) turns out to be most relevant for functions satisfying the Lipschitz condition ($\alpha = 1$).

Now we can make the following conclusion which has the status of a practical recommendation: *for functions whose order of smoothness is equal to α the kernel of Vallée-Poussin type corresponding to ρ_* given by (5.7) should be chosen in order to optimize the theoretical approximation error.* In particular, the smoother a given function the closer the kernel to the Dirichlet kernel ($\rho = 1$) and the "worse" a function the closer the kernel to the Fejér kernel ($\rho = 0$) should be. This observation is in a good correspondence to the following well-known facts. The partial sums of the Fourier series ($\rho = 1$) are most relevant for approximation of smooth functions, but a "bad" tool of approximation for "bad" functions. The Fejér means ($\rho = 0$) have the inverse property, namely, they are a "good" tool of approximation for functions with the smoothness $0 < \alpha < 1$, but they are "blind" to its improvement. It would be of interest to find similar relations for the multivariate case.

6. Some remarks on implementation

The first version of the SA-algorithm is implemented in *Maple*, which is a suitable environment for the solution of various mathematical problems. It seems to be relevant also for our procedure [3]. Our choice of a programming platform is mainly based on the following properties of *Maple*:

- possibilities for symbolic computations;
- FFT as an internal procedure;
- internal generator of pseudo-random numbers.

In this version we consider the one-dimensional case ($d = 1$), only. The function $f(x)$ as an input parameter is represented by a special procedure, so it can be given by a formula, by an algorithm or by a table of its values on the grid of equidistant nodes. However, in the last case some interpolation technique is required. In all experiments we have performed we put

$$\varepsilon = 0.15, \quad \gamma = 1.5, \quad \rho = \frac{10}{13}, \quad \sigma = 24, \quad (6.1)$$

and the function $\varphi \equiv \varphi_\rho(\xi)$ of type (5.4) has been taken as the generator of the kernel. By (5.5) the Fourier transform of φ_ρ belongs to $L_p(\mathbb{R})$, if and only if, $p > \frac{1}{2}$. In our experiments we deal with $p = 1$ and $p = \frac{3}{4}$. By (3.2) we have $n = 31, 5$ for the order of the kernel.

For the sake of convenience we rewrite formula (1.3) for nodes in the following form:

$$t_n^k = \frac{\pi k}{N}, \quad k = 0, \dots, 2N - 1 \quad \left(N = \frac{2n + 1}{2} = 32 \right). \quad (6.2)$$

As we have already mentioned above, such a form of parameters and nodes enables us to apply the FFT to calculate the Fourier-Lagrange coefficients.

In accordance with the SA-algorithm the grid is shifted by random values θ_j , $j = 1, \dots, m$, where m is given by (3.4). In our case $m = 5$ for $p = 1$ and $m = 7$ for $p = \frac{3}{4}$. By the procedure described in Section 4 we apply the standard internal Maple generator of pseudo-random numbers to construct independent and uniformly distributed random values η_j , $j = 1, \dots, m$, in the interval $[0, 1]$, so, that $\theta_j = \tau \eta_j$, where $\tau = \frac{\pi}{N}$.

In order to avoid the possible appearance of the imaginary part in the outcome, if formula (4.2) is directly applied, we use the standard real representation of the Fourier-Lagrange coefficients (θ is one of θ_j)

$$\begin{aligned} a_k &\equiv a_k^{(n)}(f; \theta) = \frac{1}{N} \sum_{\nu=0}^{2N-1} f(t_n^\nu + \theta) \cos(k(t_n^\nu + \theta)) \quad (k = 0, \dots, N), \\ b_k &\equiv b_k^{(n)}(f; \theta) = \frac{1}{N} \sum_{\nu=0}^{2N-1} f(t_n^\nu + \theta) \sin(k(t_n^\nu + \theta)) \quad (k = 1, \dots, N - 1). \end{aligned} \quad (6.3)$$

It should be noticed that applying the internal Maple FFT-procedure to the vector with the components $f(t_n^\nu + \theta)$, $\nu = 0, \dots, 2N - 1$, we obtain the values

$$\begin{aligned} c_k &= \sum_{\nu=0}^{2N-1} f(t_n^\nu + \theta) \cos(kt_n^\nu) \quad (k = 0, \dots, N), \\ d_k &= - \sum_{\nu=0}^{2N-1} f(t_n^\nu + \theta) \sin(kt_n^\nu) \quad (k = 1, \dots, N - 1), \end{aligned}$$

which do not coincide with coefficients (6.3). However, as it is easy to see, the values we need can be obtained using the formulas

$$\begin{aligned} a_k &= \frac{1}{N} (c_k \cos k\theta + d_k \sin k\theta) \quad (k = 0, \dots, N - 1), \\ b_k &= \frac{1}{N} (c_k \sin k\theta - d_k \cos k\theta) \quad (k = 1, \dots, N - 1). \end{aligned} \quad (6.4)$$

We notice that there is no need to recalculate a_n , because in accordance with formula (4.3) (see also (6.5) below) it will be multiplied by $\varphi(\frac{N}{n})$, which is equal to 0.

This way is more effective in comparison with the direct calculation by formulas (6.3) because recalculation (6.4) requires only $O(N)$ arithmetic operations; therefore, the Fourier coefficients (6.3) are calculated by $O(N \log N)$ operations.

In order to calculate the approximation polynomial given by (4.3) we rewrite this formula in its real form, that is, we use

$$\mathcal{L}_{n;\theta}^\varphi(f; x) = \frac{a_0}{2} + \sum_{k=1}^{N-1} \varphi\left(\frac{k}{n}\right) (a_k \cos kx + b_k \sin kx). \quad (6.5)$$

The approximation errors given by (4.4) are calculated by the direct rectangular method of numerical integration with the step $\frac{\pi}{100}$. Naturally, the complete procedure is repeated m times.

Let us mention that this way to compute integrals is a quite rude method, especially, if it is applied to strong oscillating functions. However, this fact practically does not influence the quality of the algorithm because of the following reasons: in contrast to the F-algorithm the SA-algorithm contains only m integrations independent of the order of the approximation polynomial; the main second step does not contain integrations.

7. Applications

In this section we describe the field of problems the SA-algorithm can be applied to. As it was already mentioned in the Introduction, the I-algorithm being a simple and economical one from the point of view of the number of operations can not be applied to non-continuous processes.

In Figure 1 the result of the approximation of the "step-function"

$$\mathcal{X}(\xi) = \begin{cases} 0, & -\pi \leq \xi < 0 \\ 2, & \xi = 0 \\ 1, & 0 < \xi < \pi. \end{cases} \quad (7.1)$$

by its interpolation mean of order 32 with the Vallee-Poussin kernel of type (5.4) with $\rho = \frac{10}{13}$ is shown ($\sigma = 24, n = 31.5$). We see that if one of the nodes coincides with the point of discontinuity the corresponding polynomial of interpolation type does not reflect the behavior of the given function.

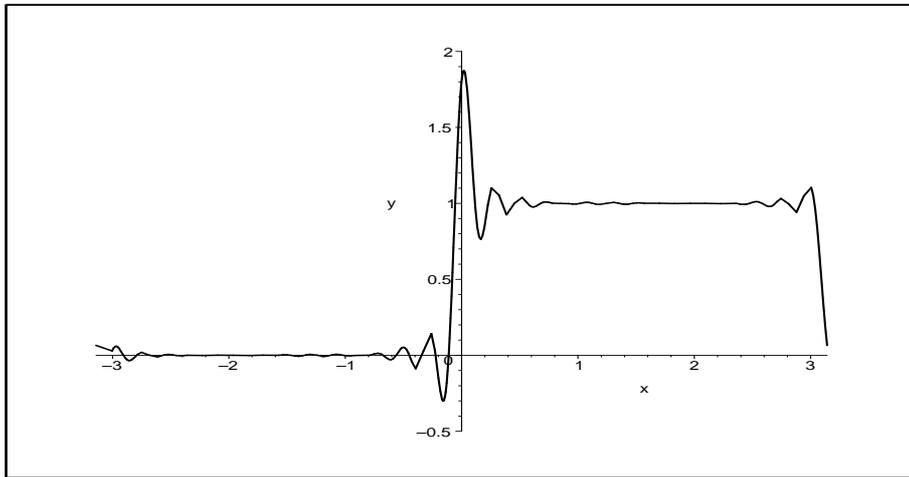


Figure 1

For continuous, but strong oscillating functions the interpolation means turn out to be not relevant as a tool of approximation as well. Indeed, for the function $\sin(\frac{n+1}{2})x$ the interpolation mean of order n is, clearly, identical 0 independently on the choice of the kernel. This observation shows that the order of an approximating polynomial should be much greater than the number of oscillations, that causes the strong increase of the computational error because of the necessity to add great numbers of different signs when calculating the Fourier-Lagrange coefficients.

In difference to the interpolation means the Fourier means can be applied to the approximation of functions in the L_p -spaces with $1 \leq p \leq +\infty$. However, such a strengthening seems to be of theoretical interest only. Indeed, the

calculation of the Fourier means is reduced to the calculation of the Fourier coefficients, that is, to integrals. There are many methods to solve this problem. One of the most popular one is the replacement of an integral by some cubature formula. Any cubature contains the values of a given function at equidistant nodes. Thus, the problem of calculating the Fourier means contains itself a problem which is quite similar to the calculation of the interpolation means. This implies that all restrictions connected with the application of the interpolation processes we discussed above are automatically valid if one deals with Fourier means.

The SA-algorithm based on approximation by families of linear polynomial operators of interpolation type does not contain the problem of calculating the integrals whose number would be comparable with the order of the approximating polynomial, and it is applicable for all L_p -spaces with $0 < p \leq +\infty$. Because of these reasons it is a relevant procedure *for approximation of discontinuous strong oscillating and even non-integrable functions*. Moreover, the number of operations turns out to be equivalent to the number of operations of the FFT.

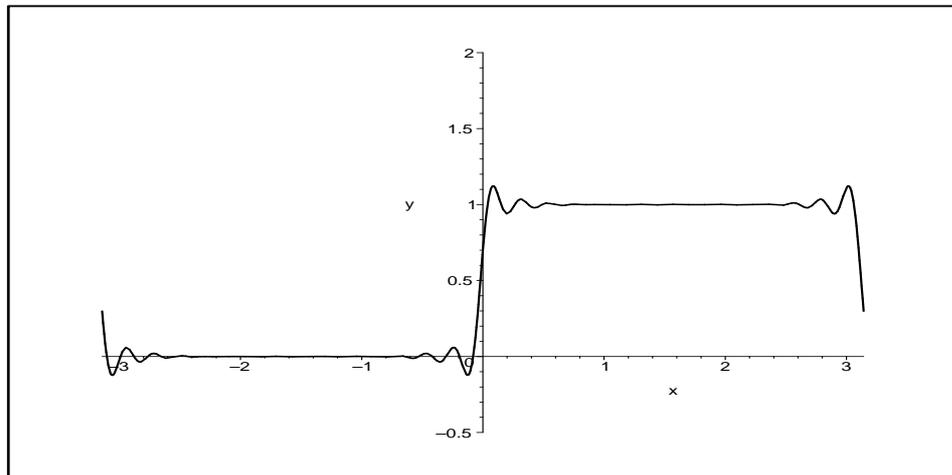


Figure 2

In the Figure 2 we can see the result of applying of the SA-algorithm to the function $\mathcal{X}(\xi)$ in the metric L_1 ($\rho = \frac{10}{13}, \sigma = 24, n = 31.5$). In this situation $m = 5$. The minimum of approximation errors in our experiment is equal to $0.2021 \dots$ which is reached, if $\eta = 0.2900 \dots$

In Figure 3 we represent the approximation of the function g given on $[0, 2\pi]$ by

$$g(x) = \begin{cases} |x|^{-1}, & x \neq 0 \\ 20, & x = 0 \end{cases}$$

in the metric $L_{3/4}$ by a relevant representative of the family of type (1.1) with the Vallée-Poussin kernel ($\rho = \frac{10}{13}, \sigma = 24, n = 31, 5$). In this case $m = 7$ and

the minimum of approximation errors in our experiment is equal to 2.9975... It corresponds to $\eta = 0.8913...$

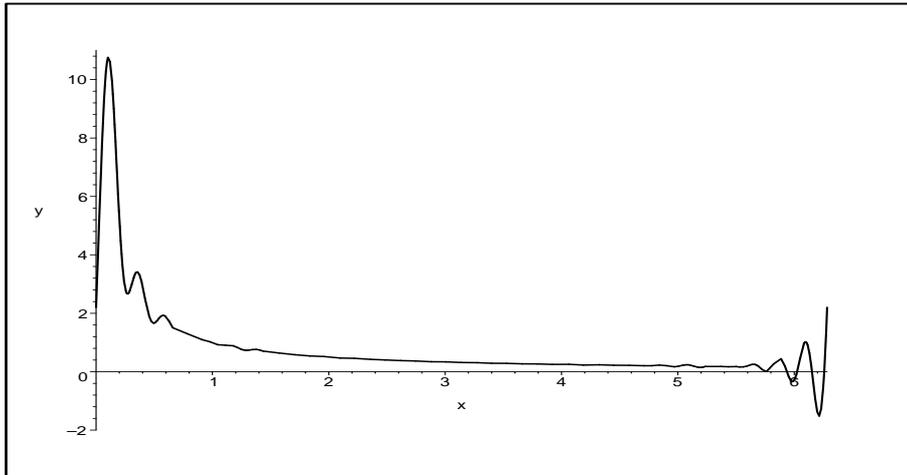


Figure 3

We come back to Figures 1 and 2. Surely, this example seems to be quite artificial because the exceptional point 0 is a priori defined. However, as it follows from our considerations, the SA-algorithm does not take into account such special data, even if it is not known where they are located. This observation enables us to apply the SA-algorithm to the processing of data containing some exceptions from the general trend which, as a rule, are caused by non-accurate measurements and many other factors out of our control, that is, being of stochastic nature. More precisely, *the SA-algorithm can eliminate the stochastic component (noise) of any determined processes concentrated on a discrete set*. If the process we measure should be continuous by some analytical argument, such a noise can be revealed "by eyes". If the process is a priori discontinuous, for instance, if it describes crises situations, then the problem of finding its main component seems to be not so easy. Recalling that the SA-algorithm is practically an interpolation procedure, but with randomly shifted nodes without loss of their "equidistancy", one can say that with the help of this procedure we "struggle" against any presence of random components in the vertical directions, which is out of our control, by means of an artificial introduction of a limited controlled random factor in the horizontal direction.

The SA-algorithm can be successfully used to *determine the nature of a local event* we are interested in. Indeed, if the approximation curve reacts on it (more precisely, on the corresponding point on the graph of data), this event, apparently, reflects the general trend in the development of the process we study and it should be taken into account for further analysis and forecasting. Otherwise, this event is accidental and will not essentially influence the further development of the process. In future work we intend to give some applications of our

method to data analysis in biology, political, economical and social sciences or medicine.

Finally, we give an example of approximation of strong oscillating functions.

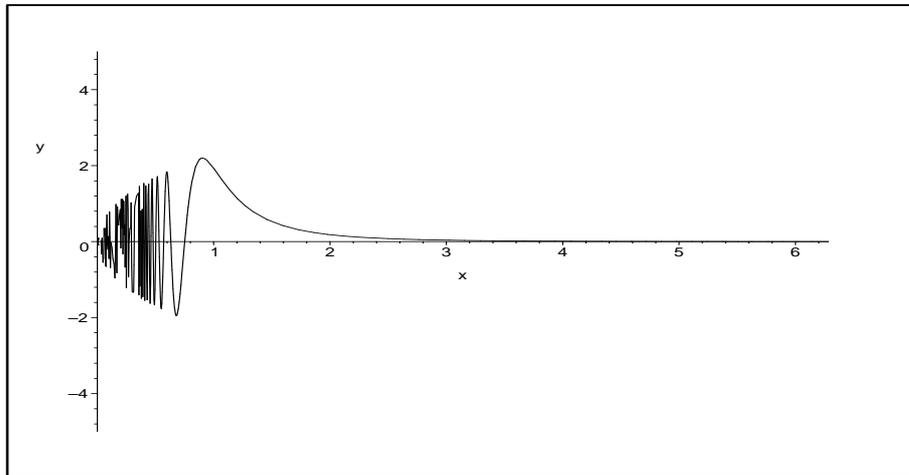


Figure 4

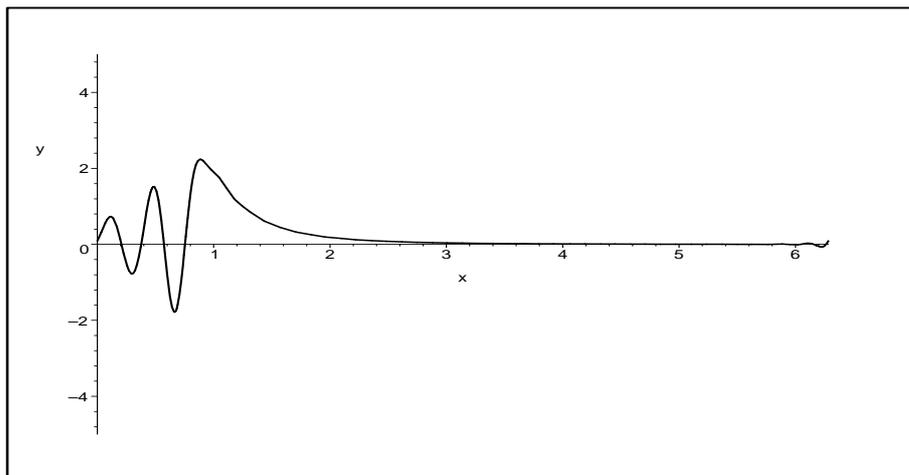


Figure 5

In Figure 4 the graph of the function

$$f(x) = \begin{cases} (x(2\pi - x))^{\frac{1}{2}} \sin \frac{1}{x^4}, & 0 < x \leq 2\pi \\ 0, & x = 0 \end{cases}$$

is shown. Clearly, the number of its oscillations is infinite. We apply the SA-algorithm to approximate f in the L_1 -metric with the parameters $\rho = \frac{10}{13}$,

$\sigma = 24, n = 31.5$. The result is represented on the Figure 5. The approximation error $0.6135\dots$ is attained for $\eta = 0.8250\dots$. It should be noticed that the performing time turned out to be less than the time one needs to calculate the first Fourier coefficient of function f by the internal Maple procedure for numerical integration. This shows that the SA-algorithm is much more economical than the algorithm based on the computation of the Fourier means.

Acknowledgement. This research was done in the framework of the AvH - project "Stochastic-analytical methods of short-time prognosis of political, social and economical processes during crises" (V - Fokoop - DEU/1025075).

References

- [1] Burinska, Z., Runovski, K. and Schmeisser, H.-J., On the method of approximation by families of linear polynomial operators. *Z. Anal. Anwendungen* 19 (2000), 677 – 693.
- [2] Cooley, J. W. and Tukey, J. W., An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation* 19 (1965), 297 – 301.
- [3] Monogan, M. B., Geddes, K. O. and Heal, K. M., *Maple V. Programming Guide*. New York: Springer 2001.
- [4] Runovski, K., On families of linear polynomial operators in L_p -spaces, $0 < p < 1$. *Russian Acad. Sci. Sb. Math.* 78 (1994), 165 – 173 (translated from *Ross. Akad. Nauk Matem. Sbornik* 184 (1993), 33 – 42).
- [5] Runovski, K., On approximation by families of linear polynomial operators in L_p - spaces, $0 < p < 1$. *Russian Acad. Sci. Sb. Math.* 82 (1995), 441 – 459 (translated from *Ross. Akad. Sci. Matem. Sbornik* 185 (1994), 81 – 102).
- [6] Lasser, R. and Runovski, K., *General convergence theory for methods of trigonometric approximation*. GSF, IBB Preprintreihe, Preprint 03-12, 2003.
- [7] Timan, A., *Theory of Approximation of Functions of a Real Variable*. New York: Dover 1994.
- [8] Walker, J. S., *Fast Fourier Transforms.*(2. ed.) Boca Raton (FL): CRC Press 1996.

Received January 17, 2005; revised February 2, 2006