Preprints, 1st IFAC Conference on Modelling, Identification and
Control of Nonlinear Systems
June 24-26, 2015. Saint Petersburg, Russia

Randomized Stochastic Approximation under Unknown but Bounded External Noise and It's Application to a Control of Educational Processes ⋆

Oleg Granichin, * **Olga Granichina**, ** **and Sergey Trapitsin** ***

* *Saint Petersburg State University (Faculty of Mathematics and Mechanics, and Research Laboratory for Analysis and Modeling of Social Processes), Institute of Problems in Mechanical Engineering, Russian Academy of Sciences, and ITMO University*
E-mail: oleg_granichin@mail.ru
** *Saint Petersburg State University (Faculty of Mathematics and Mechanics) and Herzen State Pedagogical University, St. Petersburg, Russia*
E-mail: olga_granichina@mail.ru
*** *Herzen State Pedagogical University, St. Petersburg, Russia*
E-mail: trapitsin@gmail.com

Abstract: In this paper the randomized stochastic approximation (RSA) methods for optimization problems for system unknown parameters tracking are studied. Such randomized SA procedures are working under arbitrary external observation noise. One of the possible application of the method to a control of educational processes is presented.
Keywords — randomized stochastic approximation, simultaneous perturbation, unknown but bounded external noise, control of educational processes.

1. INTRODUCTION

Recently, the development of control methods allows to find fast online solution for many practical problems. It could be done by integrating of "smart" devices in control schemes, technological processes, various decision making systems, and many others. From mathematical point of view, the optimization problems are the most important in such applications.

Stochastic approximation (SA) was introduced by Robbins and Monro [1951] and was further developed for optimization problems by Kiefer and Wolfowitz [1952] (KW-procedure). In Blum [1954] the stochastic approximation algorithm was extended to the multidimensional case. In case, when $\theta \in \mathbb{R}^d$, the conventional KW-procedure which is based on finite-difference approximations of the function gradient vector uses $2d$ observations at each iteration to construct the sequence of estimates (two observations for approximations of each component of the gradient d-vector). In late 80s, early 90s of XX century the randomized versions of SA procedure with one (or two) measurements per iteration were introduced by Granichin [1989, 1992] and Polyak and Tsybakov [1990]. Independently, similar concept was introduced by Spall [1992]. Method, introduced by Spall, is now widely referred as SPSA (*simultaneous perturbation stochastic approximation*). These algorithms recursively generate estimates along random directions. Polyak and Tsybakov [1990] proved an asymptotically optimal converge rate of algorithms. Spall [1992] demonstrated the effectiveness of algorithms in the multidimensional case (even when $d \to \infty$), and in Granichin [1989, 1992] it was established the consistency in presence of an arbitrary external noise. The possibility of application for non-stationary problems (for tracking) was studied Granichin and Amelina [2015]. General overview of methods was presented in Granichin et al. [2015].

The stochastic approximation method has a wide range of applications in different fields. RSA consistency in presence of an arbitrary external noise allows using it to analysis and in modeling of social events and processes when existence of strong (perfect) models is doubtful. Generally, one of the important possible applications is a control of educational processes. The importance of education quality objective assessment is growing in the world. Different countries would like to be sure that their educational systems are trusted, and they strive to demonstrate the ability to provide high quality education according to the international standards. Many developed countries started to be actively engaged in these issues only during the last decades. Many conceptual models, which describe the education process and factors that have potentially impact on the quality, were developed Barr and Dreeben [1983], Willms [1992], Beasley [1995], Mc Millan and Wing [2006], Katharakis and Katharaki [2010], Kempkesa and Pohl [2010]. Data Envelopment Analysis (DEA) is widely used to analyze the quality of educational systems (see,

⋆ The authors acknowledge the Russian Ministry of Education and Science (project 27.1835.2014/K) and RFBR (project 13-07-00250).

e.g., Athanassopoulos and Shale [1997], Avkiran [2001], Abbott and Doucouliagos [2003], Johnes [2006]). DEA does not require a complete specification for the functional form of the educational process, as well as the distribution of inefficient deviations. It only requires a general assumptions. However, if those assumptions are too weak, inefficiency levels may be systematically underestimated in small samples. In addition, erroneous assumptions may cause inconsistency with a bias.

The paper is organized as follows. In Section II, the problem statement is described, and the basic assumptions are made. In Section III, the main theoretical result with proof is presented. Section IV presents the problem of a control of educational processes and an application example. Simulation results are given in Section V. Section VI contains conclusion remarks.

2. MEAN-RISK OPTIMIZATION PROBLEMS AND EXCITING TESTING PERTURBATION

Consider a set of differentiable functions $\{f_w(\theta)\}_{w\in\mathbb{W}}$, $f_w(\theta):\mathbb{R}^d\to\mathbb{R}$, and let $\mathbf{x}_1,\mathbf{x}_2,\ldots$ be a set of observation points chosen by experimenter. For each $t=1,2,\ldots$ we obtain measurements $y_1,y_2,\ldots$ of $f_w(\cdot)$ with additive external noise v_t

$$y_t = f_{w_t}(\mathbf{x}_t) + v_t, \qquad (1)$$

where $\{w_t\}$ is an uncontrollable sequence, $w_t\in\mathbb{W}$.

We adopt following notations: E is a symbol for mathematical expectation, $\mathcal{F}_{t-1}$ stands for σ-algebra of all probabilistic events occurred before $t=1,2,\ldots$, $E_{\mathcal{F}_{t-1}}$ is a symbol for conditional mathematical expectation with respect to $\mathcal{F}_{t-1}$, $\langle\cdot,\cdot\rangle$ is a scalar product of two vectors.

The problem is to find "drifting" points θ_t that minimize mean-risk functionals

$$F_t(\theta) = E_{\mathcal{F}_{t-1}} f_{w_t}(\theta) \to \min_{\theta}. \qquad (2)$$

Let Δ_n, $n=1,2,\ldots$ be an observed sequence of independent Bernoulli random vectors in $\mathbb{R}^d$ with values $\pm\frac{1}{\sqrt{d}}$ with probability $\frac{1}{2}$, called the simultaneous test perturbation.

At first, we take a fixed initial vector $\widehat{\theta}_0\in\mathbb{R}^d$ and choose positive constants α and β. Consider the randomized stochastic approximation (RSA) algorithm for constructing sequences of points of observations $\{\mathbf{x}_n\}$ and estimates $\{\widehat{\theta}_n\}$

$$\begin{cases} \mathbf{x}_{2n-\frac{1}{2}\pm\frac{1}{2}} = \widehat{\theta}_{n-1} \pm \beta\Delta_n, \\ \widehat{\theta}_n = \widehat{\theta}_{n-1} - \dfrac{\alpha_n}{2\beta}\Delta_n(y_{2n}-y_{2n-1}). \end{cases} \qquad (3)$$

Next, we assume the following about $f_w(\mathbf{x})$, $F(\mathbf{x})$ and disturbances:
A1: $\forall t,\mathbf{x}\in\mathbb{R}^d$ and θ_t

$$\langle \mathbf{x}-\theta, E_{\mathcal{F}_{t-1}}\nabla_{\mathbf{z}} f_{w_t}(\mathbf{x})\rangle \geq \mu\|\mathbf{x}-\theta\|^2$$

with constant $\mu>0$.
A2: $\forall w_t\in\mathbb{W}$ gradient $\nabla f_{w_t}(\mathbf{x})$ satisfies the Lipschitz condition: $\forall\mathbf{x}',\mathbf{x}''\in\mathbb{R}^d$

$$\|\nabla f_{w_t}(\mathbf{x}') - \nabla f_{w_t}(\mathbf{x}'')\| \leq M\|\mathbf{x}'-\mathbf{x}''\|$$

with constant $M\geq\mu$.
A3: Gradient $\nabla f_{w_t}(\cdot)$ is uniformly bounded in the mean-squared sense at minimum points θ_t: $\|E\nabla\tilde{f}_t(\theta_t)\|\leq c_1$, $E\|\nabla\tilde{f}_t(\theta_t)\|^2\leq c_2$ ($c_1=c_2=0$ if w_t is not random, i.e. $f_{w_t}(\mathbf{x})=F_t(\mathbf{x})$).
A4: The drift is bounded: $\|\theta_t-\theta_{t-1}\|\leq\delta_\theta<\infty$. The rate of drift is bounded: $E_{\mathcal{F}_{2n-2}}\varphi_n(\mathbf{z})^2\leq c_3\|\mathbf{z}-h(\theta_{2n-2})\|^2+c_4$, where $\varphi_n(\mathbf{x})=f_{w_{2n}}(\mathbf{x})-f_{w_{2n-1}}(\mathbf{x})$.
A5: $\forall n$ **a)** successive differences $\bar{v}_n=v_{2n}-v_{2n-1}$ of observation noises are bounded: $|\bar{v}_n|\leq c_v<\infty$.
b) If $\bar{v}_n$ is random then $\bar{v}_n$ and vector $\mathbf{\Delta}_n$ are independent.
c) $\mathbf{\Delta}_n$ and w_{2n-1},w_{2n} (if they are random) do not depend on σ-algebra $\mathcal{F}_{2n-2}$.
d) If w_{2n-1},w_{2n} are random then random vectors $\mathbf{\Delta}_n$ and elements w_{2n-1},w_{2n} are independent.

3. MAIN RESULT

We introduce the quality characteristic of estimates sequence.

Definition. A sequence of estimates has asymptotically optimal mean upper bound $\bar{L}>0$ if $\forall\varepsilon>0\ \exists N$ such that $\forall n>N\ \sqrt{E\|\widehat{\theta}_{2n}-\theta_{2n}\|^2}\leq\bar{L}+\varepsilon$.

Denote $\gamma=3\alpha(M^2+\frac{c_3}{\beta})$, $m=2(\mu-\gamma)$, $b=\beta M(1+6\alpha M)+\delta_\theta(M+2\mu+6\alpha M^2)$, $\bar{l}=2\delta_\theta(M(2\beta+\delta_\theta)+c_1+3\mu\delta_\theta^2)+2\alpha(c_v^2+3(\frac{c_4}{\beta}+c_2+M^2(\delta_\theta+\beta))$.

The following Theorem, which extends the corresponding results from Granichin et al. [2009, 2015] and gives more accurate bound for the considered case than in Granichin and Amelina [2015], shows the asymptotically efficient mean upper bound of estimation residuals by algorithm (3).

Theorem 1. **If** assumptions **A1-A5** hold, and α is sufficiently small: $\alpha\in(0;\mu/\gamma)$ when $\mu^2>2\gamma$, or $\alpha\in(0;\frac{\mu-\sqrt{\mu^2-2\gamma}}{2\gamma})\cup(\frac{\mu+\sqrt{\mu^2-2\gamma}}{2\gamma};\mu/\gamma)$ otherwise,

then the sequence of estimates provided by algorithm (3), has asymptotically optimal mean upper bound

$$\bar{L}=(b+\sqrt{b^2+m\bar{l}})/m. \qquad (4)$$

Proof. Denote $\nu_n=\|\widehat{\theta}_{2n}-\theta_{2n}\|$, $\bar{f}_n=f_{w_{2n}}(\mathbf{x}_{2n})-f_{w_{2n-1}}(\mathbf{x}_{2n-1})$, $s_n=\frac{\alpha}{\beta_n}(\bar{f}_n+\bar{v}_n)\mathbf{\Delta_n}$, $\mathbf{d}_t=\widehat{\theta}_{2\lceil\frac{t-1}{2}\rceil}-\theta_t$, $\tilde{\mathcal{F}}_{n-1}=\sigma\{\mathcal{F}_{2n-2},w_{2n-1},w_{2n}\}$, where $\lceil\cdot\rceil$ is a ceiling function. Since $s_n=\frac{\alpha}{\beta_n}(y_{2n}-y_{2n-1})\mathbf{\Delta}_n$, according to the observation model (1) and algorithm (3), we obtain $\nu_n^2=\|\mathbf{d}_{2n}\|^2+\|s_n\|^2-2\langle\mathbf{d}_{2n},s_n\rangle$. Hence, by taking the conditional expectation over σ-algebra $\tilde{\mathcal{F}}_{n-1}$, we could bound $E_{\tilde{\mathcal{F}}_{n-1}}\nu_n^2$ as follows: $E_{\tilde{\mathcal{F}}_{n-1}}\nu_n^2\leq\|\mathbf{d}_{2n}\|^2-$

$$2\langle\mathbf{d}_{2n},\frac{\alpha}{\beta_n}E_{\tilde{\mathcal{F}}_{n-1}}\bar{f}_n\mathbf{\Delta}_n\rangle+2\frac{\alpha^2}{\beta_n^2}E_{\tilde{\mathcal{F}}_{n-1}}(\bar{v}_n^2+\bar{f}_n^2), \qquad (5)$$

since by virtue of Assumptions **A5b** we have $E_{\tilde{\mathcal{F}}_{n-1}}\bar{v}_n\mathbf{\Delta}_n=E_{\tilde{\mathcal{F}}_{n-1}}\bar{v}_nE_{\tilde{\mathcal{F}}_{n-1}}\mathbf{\Delta}_n=E_{\tilde{\mathcal{F}}_{n-1}}\bar{v}_n\cdot 0=0$.

Taylor series of $f_{w_t}(\mathbf{x}_t)$ for $t^\pm=2n-\frac{1}{2}\pm\frac{1}{2}$ gives $f_{w_{t^\pm}}(\mathbf{x}_{t^\pm})=f_{w_{t^\pm}}(\widehat{\theta}_{2n-2})\pm\langle\nabla_{t^\pm}(\rho_{t^\pm}),\beta\mathbf{\Delta}_n\rangle$, where $\rho_{t^\pm}\in(0,1)$, $\nabla_{t^\pm}(\rho_{t^\pm})=\nabla f_{t^\pm}(\widehat{\theta}_{2n-2})\pm\rho_{t^\pm}\beta\mathbf{\Delta}_n$.

Since $E_{\tilde{\mathcal{F}}_{n-1}}\varphi_n(\widehat{\theta}_{2n-2})\mathbf{\Delta}_n = \varphi_n(\widehat{\theta}_{2n-2})E_{\tilde{\mathcal{F}}_{n-1}}\mathbf{\Delta}_n = 0$, by virtue of Assumptions **A3**, for the second term in (5) we have $-2\langle \mathbf{d}_{2n}, \frac{\alpha}{\beta_n}E_{\tilde{\mathcal{F}}_{n-1}}\bar{f}_n\mathbf{\Delta}_n\rangle \leq$

$$2\frac{\alpha}{\beta}(2\beta^2 M(\nu_{n-1} + \|\theta_{2n} - \theta_{2n-1}\| + \|\theta_{2n-1} - \theta_{2n-2}\|) -$$

$$E_{\tilde{\mathcal{F}}_{n-1}}\langle \theta_{2n} - \theta_{2n-1}, \beta_n^- \nabla_{t^-}(0)\rangle + \sum_{t^\pm}\langle \mathbf{d}_{t^\pm}, \beta\nabla_{t^\pm}(0)\rangle).(6)$$

According to the first part of Assumption **A4**, we get $E_{\mathcal{F}_{2n-2}}\|h(\mathbf{d}_{2n})\|^2 \leq E_{\mathcal{F}_{2n-2}}\|h(\mathbf{d}_{2n-1})\|^2 + 2\delta_\theta\nu_{n-1} + 3\delta_\theta^2$.

By taking consequently conditional expectations over σ-algebras $\mathcal{F}_{2n-1}$ and $\mathcal{F}_{2n-2}$ and by using Assumption **A1**, we derive

$$E_{\mathcal{F}_{2n-2}} - (\langle \theta_{2n} - \theta_{2n-1}, \beta\nabla_{t^-}(0)\rangle + \sum_{t^\pm}\langle \mathbf{d}_{t^\pm}, \beta\nabla_{t^\pm}(0)\rangle) \leq$$
$$\leq E_{\mathcal{F}_{2n-2}}\beta\|\theta_{2n} - \theta_{2n-1}\|(M\|d_{t^-}\| + c_1) - \alpha\mu\|d_{t^-}\|^2 +$$
$$+\beta\mu(2\|\theta_{2n} - \theta_{2n-1}\|\|d_{t^-}\| + \|\theta_{2n} - \theta_{2n-1})\|^2.$$

Hence, we have $-E_{\mathcal{F}_{2n-2}}2\langle \mathbf{d}_{2n}), \frac{\alpha}{\beta}\bar{f}_n\mathbf{\Delta}_n\rangle \leq$

$$-2\mu\nu_{n-1}^2 + 2\frac{\alpha}{\beta}(2\beta^2 M(\nu_{n-1} + 2\delta_\theta) + \quad (7)$$
$$\beta\delta_\theta(M\nu_{n-1} + M\delta_\theta + c_1) + \beta\mu(2\delta_\theta(\nu_{n-1} + \delta_\theta) + \delta_\theta^2)).$$

Consider squared difference $\bar{f}_n^2$. By using representations $\bar{f}_n = \varphi_n(\widehat{\theta}_{2n-2}) + \beta\sum_{t^\pm}\langle \nabla f_{w_{t^\pm}}(\theta_{t^\pm}), \mathbf{\Delta}_n\rangle + \langle \nabla_{t^\pm}(\rho_{t^\pm}) - \nabla f_{w_{t^\pm}}(\theta_{t^\pm}), \mathbf{\Delta}_n\rangle$, and the symmetrical property of Δ_n distribution, and Assumptions **A2**, we get

$$E_{\tilde{\mathcal{F}}_{n-1}}\bar{f}_n^2 \leq 3(\varphi_n\widehat{\theta}_{2n-2})^2 +$$
$$+3\beta^2((\sum_{t^\pm}\nabla f_{w_{t^\pm}}(\theta_{t^\pm}))^2 + M^2(\sum_{t^\pm}(\|\mathbf{d}_{t^\pm}\| + \beta))^2).$$

By taking the conditional expectation over σ-algebra $\mathcal{F}_{2n-2}$, we obtain

$$E_{\mathcal{F}_{2n-2}}\bar{f}_n^2 \leq 3(c_3\nu_{n-1}^2 + c_4 + \beta(c_2 + M^2((\nu_{n-1} + \delta_\theta) + 2\beta)).$$

Summing up the findings bounds and by taking the conditional expectation over σ-algebra $\mathcal{F}_{2n-2}$, we derive the following $E_{\mathcal{F}_{2n-2}}\nu_n^2 \leq (1-\alpha m)\nu_{n-1}^2 + 2\alpha b\nu_{n-1} + \alpha\bar{l}$.By taking the unconditional expectation, we see that all conditions of Lemma 1 from Granichin et al. [2009] (or from Granichin and Amelina [2015]) hold for $e_n = \sqrt{E\nu_n^2}$. This completes the proof of Theorem 1.

Remark 1. Observation noise v_t in Theorem 1 could be said "almost arbitrary" since it may either be nonrandom but bounded or it may also be a realization of some stochastic process with arbitrary internal dependencies. In particular, to prove the results of this Theorem, there is no need to assume that v_t and $\mathcal{F}_{t-1}$ are not dependent.

Remark 2. The result of Theorem 1 shows that for the case without drift the asymptotical mean upper bound $\bar{L} = \sqrt{\alpha}\sqrt{\frac{c_v^2 + 3c_2 + M^2\beta^2}{\mu}} + \beta\frac{M}{\mu} + o(\sqrt{\alpha} + \beta)$ could be made infinitely small simply by choosing sufficiently small α and β under any noise level c_v. At the same time in the case of drift, the bigger drift norm δ_θ could be compensated by choosing bigger algorithm parameters α and β. This leads to a tradeoff between making algorithm parameters smaller because of noisy observations and making them bigger due to the drift of optimal points.

4. CONTROL OF EDUCATIONAL PROCESSES QUALITY

Recent researches lead to the important conclusion: the productive management of the education quality is impossible without transition from subjective descriptions of pedagogical phenomena and processes to their rigorous and objective assessments. Undoubtedly, mathematical models of educational systems cannot disclose and explain the nature of pedagogical phenomena, they are only optional, but absolutely necessary part of educational researches. Mathematical models provide a description and identification of patterns, properties and relationships of pedagogical objects and processes by constructing their images, identical to the structure and contents of real objects or processes, it allows us to represent pedagogical reality in an abstract form that is convenient for theoretical analysis, not only quantitative but also qualitative. The analysis of statistical data of monitored parameters of the educational quality is the most simple and applicable method that are is ordinary in practice. However, this method uses only a part of characteristics that could be measured. It is often used indicators which accuracy are questionable. In particular, this concerns the application of educational statistics at the macro and meso levels of management. Moreover, there are no statistics collected on an ongoing basis for many indicators. Low level of characteristics of an educational system description adversely affects the precision of descriptions of monitoring results, this causes difficulties in determining the control action. Incorrect using of indicators could specify false benchmarks that will adversely affect the educational system. The inability to give an adequate evaluation of any characteristic of process can eliminate this feature from the priority control objects, regardless of its actual significance. Increasing of the environment uncertainty and the complexity of educational institutions activities leads inevitably to more complex systems of assessment processes and results of this activities. In this situation, the means and methods of estimation that were time-tested and proven in the past for technical systems stop working (or work inefficiently) and require significant modifications, and sometimes a complete replacement. Therefore, the main task is the development of evaluation models and methods, ensuring adequate diagnosis and identification of trends in the development of such complex area as an education quality.

We assume that the current educational process state is defined by a set of d numbers $z_1, z_2, \ldots, z_d$. Mathematically, this set could be conveniently represented as a vector $\mathbf{z}$ in d-dimensional real space $\mathbb{R}^d$. In practice, the size of vector $\mathbf{z}$ could be very large. We can regard a detailed model where components of vector $\mathbf{z}$ represent all characteristics of the university and its business units (information on fixed assets, financing, material and intellectual resources, personal and academic information about the students, promising students, graduate students, faculty, staff and so on), or we can regard a model that operates with aggregated data.

Many components of current state $\mathbf{z}$ are difficult (and sometimes impossible) to determine at any given time. For example, it is not always easy to quantify the quality of the training of a student.

A mathematical model for the description in practical use implies the selection (or formation) of a set of measured data $\mathbf{y} = col(y_1, y_2, \dots, y_m) \in \mathbb{R}^m$ which values are assumed to be available to the observer at the selected time t. (Some components of vector $\mathbf{z}$ or functions of one or more components of vector $\mathbf{z}$ could be observed. They are give components of vector $\mathbf{y}$. Observation components can be determined not only through the current state variables, but also through their preceding values). Examples of observable quantities could be ongoing assessments of a student, or an average score of a group, or an average score of a student or a group for the entire period of study. Usually, size m is much smaller than d (if a detailed model with high dimensional state space is considered).

For current time instant t, let the path in the system state space from some initial state $\mathbf{z}_0$ to current $\mathbf{z}_t$ be denoted $\mathbf{z}_t(\cdot)$, $\mathbf{y}_t$ be the current vector of observations. As a rule, the formation of observation's vector values could influence not only to the system state vector values along the trajectory, but also to a various unidentified external disturbing factors, the totality of which is denoted by $\mathbf{w}_t$. Mathematically, we could write the following formula for current observations vector

$$\mathbf{y}_t = \mathrm{R}_t(\mathbf{z}_t(\cdot), \mathbf{w}_t),$$

where $\mathrm{R}_t(\cdot,\cdot)$ is a function of trajectory $\mathbf{z}_t(\cdot)$ and unknown disturbing factors $\mathbf{w}_t$ (this function may be dependent on time instant t).

The problem of restoring (estimating) of all components of state vector $\mathbf{z}_t$ or of a part of them based on observations $\mathbf{y}_t(\cdot)$ is a classical problem of filtering. In the stationary setting (without the introduction of time-varying t) this would be a typical problem of regression analysis. These problems are well known in the mathematical literature. However, most of papers are either not considered a possible uncontrolled disturbance $\mathbf{w}$ in the model, or they are to be slight. In a stochastic problem formulation, disturbance $\mathbf{w}$ are usually considered as random elements with known statistical properties. All these assumptions are valid for a sufficiently strict compliance model, especially developed for technical systems or descriptions of natural phenomena. For information models described the complex processes that determine the behavior of groups of people, mathematical results of "classical theories" often do not give a good answer. Typical features of human behavior are unpredictability, failure to follow general scheme which created in advance. For example, final assessments of a group of students received on a particular course is an important indicator of evaluating the educational process quality. But each teacher is prone to subjectivity. This may lead to a general overestimation or underestimation. It is difficult to provide the necessary conditions for correctness of statistical studies that are based on a repeatability of an experiment, because each next year the course will be read for another group of students and maybe by another teacher. This example shows the need of developing analytical methods that do not rely on strict limitations of the model and uncontrollable disturbances.

Analytical problems are often associated with a wide range of decision problems. To formalize the description of the information model a set of control actions $\mathbf{u} = col(u_1, u_2, \dots, u_l) \in \mathbb{U} \subset \mathbb{R}^l$ are included in addition to the state vectors and observations. Here l is a size of control vector, $\mathbb{U}$ is a set of control vector's possible values (usually, it is bounded).

Examples of control actions u_i may be a general plan of admission of students, plans for the professions funding levels of certain items in the budget, the task of some benchmarks of the educational process, and so on.

Assessment of the quality of decision-making assumes the introduction of some quantitative characteristics, which defined by obtained educational process $\mathbf{z}_t(\cdot)$. In the simplest case, it could be a value of type "Yes-No", characterizing the selected control strategy $\mathbf{u}_t(\cdot)$ as good or bad. More generally, the quality assessment can be multi-criteria and consist of a set of quality functionals $\Phi_i(\mathbf{z}_t(\cdot), \mathbf{u}_t(\cdot))$, $i = 1, 2, \dots, k$. For example, for the general characteristics of the educational process one could select only two criteria: first — the number of graduates, and the second — the average quality of training. In this case $k = 2$, and

$$\Phi_1(\mathbf{z}_t(\cdot), \mathbf{u}_t(\cdot)) \in [0, +\infty], \quad \Phi_2(\mathbf{z}_t(\cdot), \mathbf{u}_t(\cdot)) \in [0, 5].$$

Setting of quality functionals allows to formalize the concept of the decision purpose. In multi-criterion case $k > 1$, a lot of goals are naturally raise, and for each goal $i = 1, 2, \dots, k$ it is possible to determine one of two types:

$$\Phi_i(\mathbf{z}_t(\cdot), \mathbf{u}_t(\cdot)) \to \max\ (\min) \ \text{ or } \ \Phi_i(\mathbf{z}_t(\cdot), \mathbf{u}_t(\cdot)) \in S_i \subset \mathbb{R}.$$

In general case, the problem of relatable to each other of different criteria is very complicated, and it could only be solved at the stage of formulation of the problem by experts in the selected domain. The typical approach is to reduce the multi-criteria problem to a one-criterion by introducing the new quality functional which includes, in some forms, the original ones.

Example. As an example we consider the problem of deciding of setting a price for a semester study in a practical classes of programming. We assume that the costs of the course does not depend on the number of students, and all students study in the same single group. We assume that only one control parameter u ($l = 1$) is used. It is the price that is covered by the student for passing the course. It is natural to consider the optimization problem: select such value u that

$$\Phi_1(\mathbf{z}_t(\cdot), u) \to \max, \ \Phi_2(\mathbf{z}_t(\cdot), u) \to \max\,.$$

Simple arguments, backed up by a practical experience of teachers, show that the simultaneous implementation of both objectives could not be achieved, because the average quality of education begins to decline sharply when the number of students increases a certain point. If we consider the new quality functional that has a form of "balanced" sum of two initial functionals

$$\bar{\Phi}(\mathbf{z}_t(\cdot), u) = \alpha_1 \Phi_1(\mathbf{z}_t(\cdot), u) + \alpha_2 \Phi_2(\mathbf{z}_t(\cdot), u),$$

then its maximization is meaningless, since with a large value of the first term the second term may seek to the lowest value. It is well enough to consider the multiplication of original functionals

$$\Phi(\mathbf{z}_t(\cdot), u) = \Phi_1(\mathbf{z}_t(\cdot), u) \times (\Phi_2(\mathbf{z}_t(\cdot), u) - 2).$$

In practice, setting prices for course does not completely determine the state vector of educational process $\mathbf{z}_t(\cdot)$. At the course beginning a number of students enrolled in it depends on the price, not directly, but indirectly. Over time, the educational process is disturbed by perturbations (disease, natural disasters and so on). The optimal choice of parameter u must not and can not depend on all this diversity information. Description of optimization information models often uses mean-risk quality functionals. It is assumed in addition to the deterministic setting that the trajectory of educational process states is a realization of a random element. We denote $\mathbb{D}$ as a set of all educational process trajectories. If we assume that random elements defining the educational process trajectory have a distribution $\mathrm{P}(\cdot)$ on $\mathbb{D}$, then the mean-risk functional could be written as

$$F(u) = -\int_{\mathbb{D}} \Phi_1(\mathbf{z}_t(\cdot), u)(\Phi_2(\mathbf{z}_t(\cdot), u) - 2)\mathrm{P}(d\mathbf{z}_t(\cdot)).$$

The problem of choosing the course price is reformulated as $u = \mathrm{argmin} F(u) =?$.

Despite the seeming simplicity of the problem statement, the solution with classical mathematical methods is impossible even for the example considered above. Even brute force is not appropriate for optimization. In practice, function $F(u)$ is not calculated for all different values of u. Typically, the form of distribution $\mathrm{P}(\cdot)$ is difficult to describe in advance. Consequently, analytical form of function $F(u)$ cannot be obtained even approximately. In addition, there are difficulties with getting the observational data.

We can set some value u at the semester beginning and calculate number y_1 of students enrolled in the course and average value y_2 of final certifications at the end of training. The first observation quite accurately describes $\Phi_1(\mathbf{z}_t(\cdot), u)$ and the second one is a "noisy" assessment of the quality of training

$$y_2 = \Phi_2(\mathbf{z}_t(\cdot), u) + v,$$

where v is an observation noise, since observation y_2 includes a teacher subjectivity and so on.

For a long time this type of problems was considered as ill-posed problems, and usually the problem was not solved in this form. Oddly enough, but introducing of random control actions into a model allows to offer the reasonable solutions. A key element of the new approach is the including of randomization into measurements processes Granichin et al. [2015]. At the household level, this type of approach is well known and it is called "the method of trial and error".

5. SIMULATION

We now describe more precisely the formal mathematical model for example considered above and the possible algorithm for selection of the course fee (price). Let the educational process state space consists of two components: $z_1 \in [0, +\infty]$ is the number of students who choose the course and pay the fee, $z_2 \in [0, 5]$ is the average quality of training. We assume that the control parameter u is the price that is paid by the student for attending the course.

Suppose, that we could get the observation with noise

$$y = z_1(z_2 - 2) + v.$$

In other words, we select control parameter u and start an experiment. We wait for experiment result y. Naturally, in such problem formulation it is impossible to choose the optimal value of u after one experiment. However, if we have an opportunity to repeat the experiment several times and select various control parameters, the optimization problem becomes meaningful. In cases where observation noise v could be assumed to be a realization of independent random variables with zero-mean and finite variance, this problem is close to classical problems of mathematical statistics. However, we have already mentioned the fact that such requirements for observation noise may be excessive due to the problem nature. In this case we could use RSA algorithm (3).

Algorithm:
1) *Initialization and coefficient selection.* Set a counter index $n = 0$. Choose initial guess $\widehat{u}_0 \in \mathbb{R}$ and fairly small step-sizes $\alpha > 0$ and $\beta > 0$.
2) *Iteration* $n \to n + 1$.

a. Generate random value $\Delta_n \in \mathbb{R}$ which equals to ± 1 w.p. $\frac{1}{2}$.
b-1. Compute next input $\bar{u}_n^-$ by the rule: $\bar{u}_n^- = \widehat{u}_{n-1} - \beta\Delta_n$.
c-1. Start the process with input $\bar{u}_n^-$ and wait new value y_n^-.
b-2. Compute next input $\bar{u}_n^+$ by the rule: $\bar{u}_n^+ = \widehat{u}_{n-1} + \beta\Delta_n$.
c-2. Start the process with input $\bar{u}_n^+$ and wait new value y_n^+.
d. Calculate the quasigradient: $\hat{G} = \Delta_n \frac{y_n^+ - y_n^-}{2\beta}$.
e. Get the new estimate: $\widehat{u}_{2n} = \widehat{u}_{2n-2} - \alpha\hat{G}$.
3) Repeat Step 2.

Instead of carrying out the practical experiments we use computer simulations of possible implementations of the educational process. The practical experience suggests to choose z_1 as a realization of a random variable with an exponential distribution with parameter $\frac{a}{u}$, and z_2 as a realization of a Gaussian random variable with distribution $\mathcal{N}(b - e^{x_1 - c}, d)$ truncated into the interval $[0; 5]$. In simulation we chose: $a = 1000$, $b = 4.5$, $c = 6$, $d = 0.5$. The physical meaning of these parameters is as follows: a is the average value of money obtained for the course, b is the maximum average value of the quality of training with a minimum size of the group, c is the critical number of students in the group, above which the quality of the training begins to fall sharply, d is the variance of the random variable associated with the quality of training.
Selected simulation parameters are close to the heuristic estimates of these quantities which compiled on the basis of experience in conducting course of programming.

The values $u_0 = 20$, and $\alpha = 1$, $\beta = 5$ were chosen for the simulation.

Three typical simulation results of sequential estimation of the optimal value u_{opt} are shown in Fig. 1. One of them corresponds to "standard" measurement noise v_k which was modeled as a realization of independent random variables uniformly distributed on the interval $[-1; 1]$, and two others results correspond to the cases of unknown but bounded noise when the teacher assesses the quality of knowledge subjectively: $v = 2$ or $v = -2$ ($c_v = 2$).

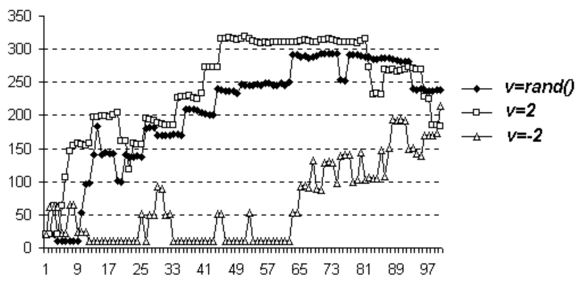

Fig. 1. Sequence of estimates $\{u_n\}$.

As could be seen, after a few tens of iterations, all three trajectories get close to the value of $u \approx 210$. The simulation has an advantage over the sequential estimation of the optimal value for the real process that we have the formal measurement model for y_n. Although, the estimation algorithm does not use this model, but for the verification of the solutions quality we are having an unlimited resource for tests which allows to calculate approximately all values of functions $F_t(u)$ for a selected type of simulation model. Fig. 2 shows the typical results of such calculations. Function $f_{w_t}(u)$ has a pronounced minimum in the neighborhood of $u = 210$. Hence the use of RSA algorithm gives a good quality assessment for different types of noise in the observations.

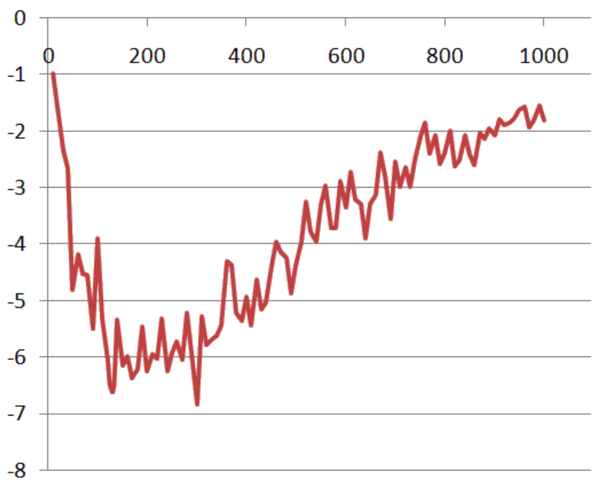

Fig. 2. The dependence of the quality functional of u.

6. CONCLUSION

In this paper the problem of tracking of minimum points of mean-risk functionals is considered. The properties of estimates of randomized stochastic approximation algorithm are studied. The application in the control of the educational processes quality is analyzed and the simulation example is discussed.

REFERENCES

M. Abbott and C. Doucouliagos. The efficiency of australian universities: A data envelopment analysis. *Economics of Education Review*, 22(1):89–93, 2003.

A.D. Athanassopoulos and E. Shale. Assessing the comparative efficiency of higher education institutions in the uk by the means of data envelopment analysis. *Education Economics*, 5(5):117–134, 1997.

N.K. Avkiran. Investigating technical and scale efficiencies of australian universities through data envelopment analysis. *Socio-Economic Planning Sciences*, 35(1):57–80, 2001.

R. Barr and R. Dreeben. *How Schools Work*. Chicago: University of Chicago, 1983.

J. Beasley. Determining teaching and research efficiencies. *Journal of the Operational Research Society*, 46(4):441–452, 1995.

J.R. Blum. Multidimensional stochastic approximation methods. *The Annals of Mathematical Statistics*, 25(4): 737–744, 1954.

O. Granichin and N. Amelina. Simultaneous perturbation stochastic approximation for tracking under unknown but bounded disturbances. *IEEE Trans. Automat. Control*, 60(5), 2015.

O. Granichin, L. Gurevich, and A. Vakhitov. Discrete-time minimum tracking based on stochastic approximation algorithm with randomized differences. In *Proc. of the 48th IEEE Conference on Decision and Control (CDC2009)*, pages 5763–5767, 2009.

O. Granichin, Z. Volkovich, and D. Toledano-Kitai. *Randomized Algorithms in Automatic Control and Data Mining*. Springer, 2015.

O.N. Granichin. A stochastic recursive procedure with correlated noise in the observation, that employs trial perturbations at the input. *Vestnik Leningrad University: Math*, 22(1):27–31, 1989.

O.N. Granichin. Unknown function minimum point estimation under dependent noise. *Problems of Information Transmission*, 28(2):16–20, 1992.

J. Johnes. Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economics of Education Review*, 25(3):273–288, 2006.

G. Katharakis and M. Katharaki. A comparative assessment of greek universities' efficiency using quantitative analysis. *International Journal of Educational Research*, 49(4–5):115–128, 2010.

G. Kempkesa and C. Pohl. The efficiency of german universities — some evidence from non parametric and parametric methods. *Applied Economics*, 42(16):2063–2079, 2010.

J. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *Ann. Math. Stat.*, 23 (3):462–466, 1952.

M. Mc Millan and H.C. Wing. University efficiency: A comparison and consolidation of results from stochastic and non-stochastic methods. *Education Economics*, 14 (1):1–30, 2006.

B.T. Polyak and A.B. Tsybakov. Optimal order of accuracy of search algorithms in stochastic optimization. *Problems of Information Transmission*, 26(2):45–53, 1990.

H. Robbins and S. Monro. A stochastic approximation method. *Ann. Math. Stat.*, pages 400–407, 1951.

J.C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3):332–341, 1992.

J. D. Willms. *Monitoring School Performance: A Guide for Educators*. Washington, DC: Falmer, 1992.