

УДК 519.21

MSC 90C15, 60G10, 60F10

## ON LARGE DEVIATIONS FOR THE SOLUTIONS OF STOCHASTIC PROGRAMMING PROBLEMS

P. S. КНОПОВ, E. J. KASITSKAYA

V. M. Glushkov Institute of Cybernetics, the National Academy of Sciences of Ukraine, Kiev, Ukraine, E-mail: knopov1@yahoo.com, e.kasitskaya@gmail.com.

## ПРО ВЕЛИКІ ВІДХИЛЕННЯ ДЛЯ РОЗВ'ЯЗКІВ ЗАДАЧ СТОХАСТИЧНОГО ПРОГРАМУВАННЯ

П. С. КНОПОВ, Є. Й. КАСИЦЬКА

Інститут кібернетики імені В. М. Глушкова, Національна академія наук України, Київ, Україна, E-mail: knopov1@yahoo.com, e.kasitskaya@gmail.com.

**ABSTRACT.** The work deals with stochastic programming problems for stationary random sequences, stationary processes, homogeneous random fields with discrete and continuous parameters. Trajectories of processes and fields are continuous. Stationary and non-stationary observations of processes and fields are considered. The former criterion function is approximated by the empirical one. It is assumed that the first problem has a unique solution. Consistency of empirical estimates for non-stationary observations is proved. Borel–Cantelli lemma is used for proving. Processes and fields are assumed to satisfy the strong mixing condition. Some restrictions on the moments of processes and fields must be fulfilled. Large deviations of the solutions are estimated. For proving the results theorems from functional analysis and large deviations theory are used. Additional conditions on behavior of minimizing function in the neighborhood of the minimum point are supposed. No stationary model is considered for the convex criterion function. The processes and fields need to satisfy the first hypothesis of hypermixing.

**KEYWORDS:** stochastic programming problem, a stationary process, a homogeneous field, large deviations.

**АНОТАЦІЯ.** Розглядаються задачі стохастичного програмування для стаціонарних випадкових послідовностей, стаціонарних процесів, однорідних випадкових полів з дискретними та неперервними параметрами. Траєкторії процесів та полів неперервні. Досліджуються стаціонарні та нестаціонарні спостереження процесів та полів. Первинна функція критерію апроксимується емпіричною. Вважається, що первинна задача має єдиний розв'язок. Доводиться консистентність емпіричних оцінок для нестаціонарних спостережень. Для доведення використовується лема Бореля–Кантеллі. Робиться припущення, що процеси та поля задовольняють умові сильного перемішування. Деякі обмеження

накладаються на моменти процесів та полів. Оцінюються великі відхилення розв'язків. Для доведення результатів використовуються теореми з функціонального аналізу та теорії великих відхилень. Накладаються додаткові умови на поведінку мінімізуємої функції в околі точки мінімуму. Нестационарна модель розглядається для опуклої функції критерію. Робиться припущення, що процеси та поля задовольняють першій гіпотезі гіперперемішування.

КЛЮЧОВІ СЛОВА: задача стохастичного програмування, стаціонарний процес, однорідне поле, великі відхилення.

#### INTRODUCTION

When one needs make a decision in uncertainty conditions, a stochastic programming problem occurs. The mean of a factor of management quality depending from a random parameter is optimized.

Indirect methods for solving of stochastic programming problems consist in approximation of a stochastic problem by associated deterministic one. One of basic indirect methods is so called empirical means method when the factors are approximated by their empirical estimates.

The main problem is estimating of precision and investigation of convergence for such approximation when quantity of observations increases.

#### 1. STOCHASTIC OPTIMIZATION PROBLEM FOR A RANDOM SEQUENCE

Let us consider the next model. One has the stochastic optimization problem

$$\min F(x) = Ef(x) = Ef(x, \xi_0), \quad x \in X, \quad (1)$$

where  $\{\xi_i, i \in \mathbb{Z}\}$  is a stationary in a strict sense metrically transitive random sequence defined on a probabilistic space  $(\Omega, G, P)$  with values in some measurable space  $(Y, \mathfrak{F})$ ;  $X$  is a nonempty compact subset of  $\mathbb{R}$ ,  $f : X \times Y \rightarrow \mathbb{R}$  is some known function continuous on the first argument and measurable on the second one.

Replace (1) by the empirical function

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n f(x, \xi_i), \quad x \in X, \quad (2)$$

where  $\{\xi_i, i = 1, \dots, n\}$  are observed elements of the sequence  $\{\xi_i\}$ .

Under some restrictions on the first moment of the function  $f$  ([1]) there exists a solution  $x_0$  for the problem (1). Suppose that it is unique.

It is known that there exists at least one minimum point  $x_n(\omega)$  for the function (2), which is a measurable function of  $\omega$ . Under some sufficiently no constricting conditions ([1])  $x_n(\omega)$  converges to  $x_0$  with probability 1 when  $n \rightarrow \infty$ .

Let us investigate large deviations for  $x_n$  and  $F_n(x_n)$  from  $x_0, F(x_0)$ . Suppose that for all  $y \in Y$  one has

$$f(\circ, y) - Ef(\circ) \in K$$

where  $K$  is some convex compact subset of  $C(X)$ . Denote

$$A_\varepsilon = \{z \in K : \|z\| \geq \varepsilon\},$$

$$\Lambda(Q) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left( E \left\{ \exp \left( \sum_{i=1}^n \int_X [f(x, \xi_i) - F(x)] Q(dx) \right) \right\} \right),$$

$$I(z) = \Lambda^*(z) = \sup \left\{ \int_X z(x) Q(dx) - \Lambda(Q), \quad Q \in M(X) \right\}.$$

**Theorem 1.** *When  $\{\xi_i\}$  satisfies the first hypothesis of hypermixing ([2])*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P \left\{ \left| \min_{x \in X} F_n(x) - \min_{x \in X} F(x) \right| \geq \varepsilon \right\} \leq -\inf \{ I(z), z \in A_\varepsilon \}. \quad (3)$$

*Suppose that there exists a conditioning function  $\psi$  for  $F(\circ)$  at  $x_0$  with some constant  $\rho$  ([3]). Let  $x_n$  be a minimum point of (2) on the set  $B(x_0, \rho)$ . If  $\varepsilon$  is so small that*

$$\psi(|x - x_0|) \leq 2\varepsilon \implies |x - x_0| \leq \rho,$$

*then one has*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P \{ \psi(|x - x_0|) \geq 2\varepsilon \} \leq -\inf \{ I(z), z \in A_\varepsilon \}. \quad (4)$$

*More if  $\psi$  is convex and strictly increasing on  $[0, \rho]$ , then*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P \{ (|x_n - x_0|) \geq \psi^{-1}(2\varepsilon) \} \leq -\inf \{ I(z), z \in A_\varepsilon \}. \quad (5)$$

*Proof.* As concluded in [3] for each  $\omega$

$$| \min \{ Ef(x), x \in X \} - \min \{ F_n(x), x \in X \} | \leq \|F_n - Ef\|. \quad (6)$$

Then, the conditions from [4] are fulfilled for the sequence  $\{\xi_i\}$ . Therefore for any  $\varepsilon > 0$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P \{ \|F_n - Ef\| \geq \varepsilon \} \leq -\inf \{ I(z), z \in A_\varepsilon \}. \quad (7)$$

The inequality (3) follows from (6) and (7).

To prove the second part of the theorem we also use [4]. Under the conditions of the theorem we have for all  $\omega$

$$\psi(|x^* - x_n|) \leq 2\|F_n - Ef\|, \quad (8)$$

or

$$|x^* - x_n| \leq \psi^{-1}(2\|F_n - Ef\|). \quad (9)$$

Taking into account (7), the inequalities (8) and (9) imply (4) and (5) respectively. The theorem is proved.  $\square$

2. STOCHASTIC OPTIMIZATION PROBLEM WITH NO STATIONARY  
OBSERVATIONS

Let us consider the next model where observations are nonstationary.

Suppose that  $\{\xi_i, i \in \mathbb{Z}\}$  is a stationary in a strict sense ergodic random sequence defined on a full probabilistic space  $(\Omega, G, P)$ , with values in some metric space  $(Y, \rho)$ ;  $X = [a; b] \subset \mathbb{R}$ ;  $h : \mathbb{R} \times X \times Y \rightarrow \mathbb{R}$  is a continuous function convex on the second argument.

Investigate the problem

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n h(i, x, \xi_i) \rightarrow \min, \quad x \in X. \quad (10)$$

Let the next conditions be fulfilled:

1.  $\sup\{E[\max |h(i, x, \xi_i)|, x \in X], i \in \mathbb{N}\} < \infty$ ;
2. for any  $x \in X$  there exists

$$F(x) = \lim_{n \rightarrow \infty} EF_n(x);$$

3. there exist such  $x_0 \in X, c > 0$ , that

$$F(x) \geq F(x_0) + c|x - x_0|, \quad x \in X. \quad (11)$$

By (11) one has that  $x_0$  is a unique solution of the problem

$$F(x) \rightarrow \min, \quad x \in X. \quad (12)$$

**Theorem 2.** *Let the sequence  $\{\xi_i\}$  satisfy strong mixing conditions with the coefficient*

$$\alpha(\tau) \leq \frac{c_0}{1 + \tau^{1+\varepsilon}}, \quad \varepsilon > 0.$$

*Suppose that expectations of right and left derivatives of the function  $F_n$  in  $x_0$  converge to such derivatives of the function  $F$ . Assume also the next conditions to be fulfilled:*

1. *it exists such  $\delta > \frac{2}{\varepsilon}$ , that for any  $i \in \mathbb{N}$  one has*

$$E|h'_+(i, x_0, \xi_i)|^{2+\delta} < \infty, \quad E|h'_-(i, x_0, \xi_i)|^{2+\delta} < \infty;$$

2. *there exists such  $c' > 0$ , that for all  $i \in \mathbb{N}$*

$$E[h'_+(i, x_0, \xi_i)]^2 < c', \quad E[h'_-(i, x_0, \xi_i)]^2 < c'.$$

*Then it exists with probability 1 such  $n_0 = n_0(\omega)$ , that for all  $n > n_0$  the problem (10) has a unique solution  $x_n = x_0$ .*

*Proof.* Denote

$$\theta_n = (F_n)'_+(x_0) - E\{(F_n)'_+(x_0)\}.$$

Then

$$\begin{aligned} E\theta_n^2 &= E\left[\frac{1}{n} \sum_{i=1}^n h'_+(i, x_0, \xi_i) - E\frac{1}{n} \sum_{i=1}^n h'_+(i, x_0, \xi_i)\right]^2 = \\ &= E\left[\frac{1}{n} \sum_{i=1}^n h'_+(i, x_0, \xi_i) - \frac{1}{n} \sum_{i=1}^n E h'_+(i, x_0, \xi_i)\right]^2 = \end{aligned}$$

$$\begin{aligned}
 &= E \left[ \frac{1}{n} \sum_{i=1}^n [h'_+(i, x_0, \xi_i) - E h'_+(i, x_0, \xi_i)] \right]^2 = \\
 &= E \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n [h'_+(i, x_0, \xi_i) - E h'_+(i, x_0, \xi_i)] [h'_+(k, x_0, \xi_k) - E h'_+(k, x_0, \xi_k)] = \\
 &= E \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n \beta_i \beta_k,
 \end{aligned}$$

where  $\beta_i = h'_+(i, x_0, \xi_i) - E h'_+(i, x_0, \xi_i)$ .

One has

$$|E \beta_i \beta_k| \leq c_1 / (1 + |i - k|^{1+\varepsilon'}).$$

Consequently

$$E \theta_n^2 \leq \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n \frac{c_1}{1 + |i - k|^{1+\varepsilon'}} \leq \frac{c_2}{n}.$$

Because of Borel–Cantelly lemma

$$P \{ \theta_{n^2} \rightarrow 0, n \rightarrow \infty \} = 1.$$

Assume

$$\varphi_m = \sup |\theta_n - \theta_{m^2}|, \quad m^2 \leq n \leq (m+1)^2.$$

Then

$$|\theta_n| \leq |\theta_{m^2}| + \varphi_m.$$

One has

$$\begin{aligned}
 \theta_n - \theta_{m^2} &= \frac{1}{n} \sum_{i=1}^n \beta_i - \frac{1}{m^2} \sum_{i=1}^{m^2} \beta_i = \\
 &= \frac{1}{n} \sum_{i=1}^n \beta_i - \frac{1}{n} \sum_{i=1}^{m^2} \beta_i + \frac{1}{n} \sum_{i=1}^{m^2} \beta_i - \frac{1}{m^2} \sum_{i=1}^{m^2} \beta_i = \\
 &= \frac{1}{n} \sum_{i=m^2+1}^n \beta_i + \left( \frac{m^2}{n} - 1 \right) \theta_{m^2}.
 \end{aligned}$$

Denote

$$\delta_m = \sup \left| \frac{1}{n} \sum_{i=m^2+1}^n \beta_i \right|, \quad m^2 \leq n \leq (m+1)^2.$$

Consequently

$$\varphi_m \leq \delta_m + \sup \left| \left( \frac{m^2}{n} - 1 \right) \theta_{m^2} \right|, \quad m^2 \leq n \leq (m+1)^2. \quad (13)$$

Then

$$E \delta_m^2 = E \sup \frac{1}{n^2} \sum_{i=m^2+1}^n \sum_{k=m^2+1}^n \beta_i \beta_k, \quad m^2 \leq n \leq (m+1)^2 \leq$$

$$\leq E \frac{1}{m^4} \sum_{i=m^2+1}^{(m+1)^2} \sum_{k=m^2+1}^{(m+1)^2} |\beta_i \beta_k| \leq \frac{c_3}{m^4} [(m+1)^2 - m^2]^2 \leq \frac{c_4}{m^2}.$$

One has

$$P \{ \delta_n \rightarrow 0, n \rightarrow \infty \} = 1.$$

For  $m^2 \leq n \leq (m+1)^2$

$$\left| \frac{m^2}{n} - 1 \right| \leq \left| \frac{m^2}{(m+1)^2} - 1 \right| = \left| \frac{1}{\left(\frac{m+1}{m}\right)^2} - 1 \right| \leq 1,$$

so (13) implies

$$P \{ \varphi_n \rightarrow 0, n \rightarrow \infty \} = 1.$$

Then

$$P \{ \theta_n \rightarrow 0, n \rightarrow \infty \} = 1.$$

Analogous relations take place for a left derivative.

Because of (11)

$$F'_+(x_0) \geq c, \quad F'_-(x_0) \geq c.$$

Then with probability 1 beginning from some  $n_0$

$$(F_n)'_+(x_0) > 0, \quad (F_n)'_-(x_0) > 0. \quad (14)$$

Convexity of  $F_n$  and (14) imply the theorem.  $\square$

**Theorem 3.** *Let the sequence  $\{\xi_i, i \in \mathbb{N}\}$  satisfy the first hypermixing hypothesis. Suppose that the function  $h$  does not depend from  $i$ , and there exists such  $L > 0$  that for all  $i, \omega$*

$$|h'_+(x_0, \xi_i)| \leq L, \quad |h'_-(x_0, \xi_i)| \leq L.$$

Then

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P(A_n^c) \leq -\inf \{V^*(z), z \in [-L; 0]\}, \quad (15)$$

where  $V^*(z) = \sup \{zQ(X) - V(Q), Q \in M(X)\}$ ,

$$V(Q) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln E \exp \left\{ Q(X) \sum_{i=1}^n \min [h'_+(x_0, \xi_i), h'_-(x_0, \xi_i)] \right\},$$

$$A_n = \{\omega : \operatorname{argmin} F_n(x) = \{x_0\}, x \in X\}, \quad A_n^c = \Omega - A_n.$$

*Proof.* One has

$$\begin{aligned} P(A_n^c) &= P \{ \min [(F_n)'_+(x_0), (F_n)'_-(x_0)] \in [-L; 0] \} \leq \\ &\leq P \left\{ \frac{1}{n} \sum_{i=1}^n \min [h'_+(x_0, \xi_i), h'_-(x_0, \xi_i)] \in [-L; 0] \right\}. \end{aligned} \quad (16)$$

Denote

$$K = \{\alpha(x) = \alpha, x \in X; \alpha \in [-L; L]\}.$$

It is evident that  $K$  is a compact convex subset of  $C(X)$ .

Let us investigate the function

$$a_i = a_i(x) = \min [h'_+(x_0, \xi_i), h'_-(x_0, \xi_i)], \quad x \in X.$$

For all  $i, \omega$  one has  $a_i(\bullet) \in K$ . Assume

$$K_1 = \{\alpha(x) = \alpha, x \in X; \alpha \in [-L; 0]\}.$$

Evidently,  $K_1$  is a closed subset of  $K$ . Then

$$\begin{aligned} P \left\{ \frac{1}{n} \sum_{i=1}^n \min [h'_+(x_0, \xi_i), h'_-(x_0, \xi_i)] \in [-L; 0] \right\} = \\ = P \left\{ \left( \frac{1}{n} \sum_{i=1}^n a_i = \frac{1}{n} \sum_{i=1}^n a_i(x), x \in X \right) \in K_1 \right\}. \end{aligned} \quad (17)$$

Because of [4] one has

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P \left\{ \frac{1}{n} \sum_{i=1}^n a_i \in K_1 \right\} \leq -\inf \{V^*(z), z \in K_1\}, \quad (18)$$

where  $V^*(z) = \sup \{zQ(X) - V(Q), Q \in M(X)\}$ ,

$$V(Q) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln E \exp \left\{ \sum_{i=1}^n \int_X a_i Q(dx) \right\} = \lim_{n \rightarrow \infty} \frac{1}{n} \ln E \exp \left\{ Q(X) \sum_{i=1}^n a_i \right\}.$$

Relations (16)–(18) imply (15).  $\square$

Analogous results take place for a homogeneous in a strict sense random field with a discrete parameter, but it is supposed that the field satisfies the strong mixing condition with the corresponding coefficient, and the moments of the field satisfy some conditions.

### 3. STOCHASTIC PROGRAMMING PROBLEMS FOR CONTINUOUS PARAMETERS

Such theorems are also right for a stationary in a strict sense random process with continuous time and continuous trajectories when the problem

$$\min F(x) = Ef(x, \xi(0)), \quad x \in X, \quad (19)$$

is approximated by the problem

$$\min F_T(x) = \frac{1}{T} \int_0^T f(x, \xi(t)) dt, \quad x \in X, \quad (20)$$

where  $\{\xi(t), t \in [0, T]\}$  are observations of the process  $\{\xi(t)\}; T > 0; f$  is a continuous function.

When observations are not stationary for continuous variant one has the empirical function

$$F_T(x) = \frac{1}{T} \int_0^T h(t, x, \xi(t)) dt, \quad x \in X,$$

where  $h$  is a continuous function convex on the second argument.

Analogous results are right for homogeneous random fields with a continuous parameter. There one has the empirical functions

$$F_{T_1 T_2}(x) = \frac{1}{T_1 T_2} \int_0^{T_1} \int_0^{T_2} f(x, \xi(t_1, t_2)) dt_1 dt_2, \quad x \in X, \quad T_1, T_2 > 0;$$

$$F_{T_1 T_2}(x) = \frac{1}{T_1 T_2} \int_0^{T_1} \int_0^{T_2} h(t_1, t_2, x, \xi(t_1, t_2)) dt_1 dt_2,$$

where the function  $h$  is continuous on all arguments and convex on  $x$ .

#### FINAL CONCLUSIONS

In consumption it should be noted that the results can be used for solving various stochastic optimization problems from recognition theory, regression analysis where one needs to find optimal decisions for observations of a stationary random process or a homogeneous random field.

The work was performed with the partial support of the National Research Fund of Ukraine (grant №2020.02/0121 "Analytical methods and machine learning in control theory and decision-making in conditions of conflict and uncertainty").

#### REFERENCES

1. Knopov P. S., Kasitskaya E. J. Properties of empirical estimates in stochastic optimization and identification problems. *Annals of Operations Research*. 1995. Vol.56. P. 225–239.
2. Deuschel J. D., Stroock D. W. Large deviations. Boston, etc.: Academ. Press, inc., 1989. 310 p.
3. Kaniovski Yu. M., King A. J., Wets R. J-B. Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems. *Annals of Operations Research*. 1995. Vol.56. P. 189–208.
4. Knopov P. S., Kasitskaya E. J. On large deviations of empirical estimates in a stochastic programming problem with no stationary observations. *Cybernetics and System Analysis*. 2010. V.46, №5. P. 46–50. (in Russian)

Received: 15.12.2021 / Accepted: 20.12.2021

### О БОЛЬШИХ УКЛОНЕНИЯХ РЕШЕНИЙ ЗАДАЧ СТОХАСТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

П. С. КНОПОВ, Е. И. КАСИЦКАЯ

Институт кибернетики имени В.М. Глушкова, Национальная академия наук Украины, Киев, Украина, E-mail: knopov1@yahoo.com, e.kasitskaya@gmail.com

**АННОТАЦИЯ.** Рассматриваются задачи стохастического программирования для стационарных случайных последовательностей, стационарных процессов, однородных случайных полей с дискретными и непрерывными параметрами. Траектории процессов и полей непрерывны. Исследуются стационарные и нестационарные наблюдения процессов и полей. Первичная функция критерия аппроксимируется эмпирической. Предполагается, что первичная задача имеет единственное решение. Доказывается состоятельность эмпирических оценок для нестационарных наблюдений. Для доказательства используется лемма Бореля Кантелли. Предполагается, что процессы и поля удовлетворяют условию

сильного перемешивания. Некоторые ограничения накладываются на моменты процессов и полей. Оцениваются большие отклонения решений. Для доказательства результатов используются теоремы из функционального анализа и теории больших отклонений. Налагаются дополнительные условия на поведение минимизируемой функции в окрестности точки минимума. Нестационарная модель рассматривается для выпуклой функции критерия. Предполагается, что процессы и поля удовлетворяют первой гипотезе гиперперемешивания.

**КЛЮЧЕВЫЕ СЛОВА:** задача стохастического программирования, стационарный процесс, однородное поле, большие отклонения.