УДК 621.391

П.С. КНОПОВ, В.И. НОРКИН

ОБ УСЛОВИЯХ СХОДИМОСТИ МЕТОДА ЭМПИРИЧЕСКИХ СРЕДНИХ В СТОХАСТИЧЕСКОМ ПРОГРАММИРОВАНИИ

Аннотация. Проанализированы условия сходимости метода эмпирических средних в стохастическом программировании при нетрадиционных условиях, когда используются зависимые наблюдения случайных параметров задачи и случайные показатели оптимизации могут быть разрывными индикаторными функциями. Для случая зависимых наблюдений установлены теоремы о вероятностях больших уклонений для приближенных оптимальных значений и решений.

Ключевые слова: стохастическое программирование, метод эмпирических средних, зависимые наблюдения, условия перемешивания, большие уклонения, разрывные функции, функции вероятности, сходимость метода.

ВВЕДЕНИЕ

Теория стохастического программирования является методологической основой принятия решений в условиях стохастической неопределенности и риска. Традиционно в задачах стохастического программирования оптимизируется среднее значение показателя качества управления, зависящего от случайного параметра [1–7].

Существуют два основных подхода к решению задач стохастического программирования: так называемые прямые стохастические итерационные методы [6, 8–10]; непрямые методы, состоящие в аппроксимации стохастической задачи приближенной детерминированной задачей [2, 3, 5]. Прямые методы незаменимы в тех случаях, когда нет информации об аналитическом виде показателей качества управления стохастической системой, а доступны только их случайные оценки или оценки чувствительности показателей к изменению параметров управления. Непрямые методы применимы тогда, когда известна структура показателей качества оптимизации или управления стохастической системой, т.е. функциональная зависимость показателей от детерминированных и случайных параметров. Сравнение этих двух подходов приведено в работе [11].

Одним из основных непрямых методов стохастического программирования является так называемый метод эмпирических средних, который также часто называется статистическим методом, методом выборочной аппроксимации (sample average approximation) или выборочным методом Монте-Карло (Monte Carlo sampling method) [5, 12, 13]. В данном методе распределение случайных параметров задачи аппроксимируется эмпирическим распределением, а показатели качества управления аппроксимируются их эмпирическими оценками. Поэтому одной из основных проблем является оценка точности и обоснование сходимости такой аппроксимации при увеличении числа наблюдений для различных классов задач стохастического программирования.

В настоящей работе анализируется сходимость метода эмпирических средних в нетрадиционных условиях, а именно: в условиях зависимых эмпирических наблюдений; в случае разрывных показателей оптимизации. Основными инструментами исследования являются теоремы о больших уклонениях для зависимых наблюдений, оценки радемахеровских средних для семейств функций и равномерный закон больших чисел для случайных функций и отображений.

БОЛЬШИЕ УКЛОНЕНИЯ В МЕТОДЕ ЭМПИРИЧЕСКИХ СРЕДНИХ ПРИ ЗАВИСИМЫХ НАБЛЮДЕНИЯХ

Рассмотрим задачу стохастического программирования: минимизировать функцию

$$F(x) = \mathrm{E} f(x) = \mathrm{E} f(x, \xi_0), \ x \in X, \tag{1}$$

где X — непустое компактное подмножество в $\mathbb{R}^d$; $\{\xi_i, i \in \mathbb{N}\}$ — стационарная в узком смысле эргодическая случайная последовательность, заданная на вероятностном пространстве (Ω, F, P) и принимающая значения в некотором измеримом пространстве $(Y, \mathfrak{F})$; $f: X \times Y \to \mathbb{R}$ — заданная функция, непрерывная по первому аргументу и измеримая по второму; символ E — математическое ожидание по мере P; $\mathbb{R}^d$ — d-мерное евклидово арифметическое пространство; $\mathbb{N}$ — целые положительные числа.

Задача (1) аппроксимируется следующей: найти точки минимума и минимальное значение функции

$$F_n(x) = \mathrm{E}^n f(x) = \frac{1}{n} \sum_{k=1}^{n} f(x, \xi_k), \quad x \in X, \tag{2}$$

где $\{\xi_k, k = 1, \dots, n\}$ — наблюдаемые элементы последовательности $\{\xi_i\}$.

В силу свойств функции f существует измеримая функция $x_n(\omega)$ такая, что при любых n, ω число $x_n(\omega)$ является точкой минимума функции (2).

Предположим, что

$$\mathrm{E}\{\max(|f(x, \xi_0)|, x \in X)\} < \infty.$$

Тогда функция $F(x)$ непрерывна и имеет по крайней мере одну точку минимума x^*. Предположим, что эта точка единственна.

Множество работ посвящено исследованию сходимости $x_n(\omega)$ к x^* и $F_n(x_n)$ к $F(x^*)$ в некотором вероятностном смысле при $n \to \infty$, а также нахождению условий асимптотической нормальности $\sqrt{n}(x_n - x^*)$ и $\sqrt{n}(F_n(x_n) - F(x^*))$ [4, 5, 9, 12].

В частности, при выполнении приведенных выше условий имеем [13]

$$P\{x_n \to x^*, \ n \to \infty\} = 1, \ P\{F_n(x_n) \to F(x^*), \ n \to \infty\} = 1.$$

Далее рассматривается задача о нахождении вероятности больших уклонений x_n от x^* и $F_n(x_n)$ от $F(x^*)$ и приводятся условия, при которых вероятность таких уклонений экспоненциально убывает при $n \to \infty$. Для этого используются известные теоремы из теории больших уклонений и функционального анализа.

Перейдем к предположениям и изложению результатов.

Полагаем, что при каждом $y \in Y$ функция $f(\circ, y)$ является элементом пространства $C(X)$ непрерывных на $X \subset \mathbb{R}$ действительных функций. Предположим также, что существует выпуклое компактное множество $K \subset C(X)$ такое, что для любого $y \in Y$ имеем $f(\circ, y) - \mathrm{E} f(\circ) \in K$. Тогда при всех ω в силу выпуклости K имеет место

$$\mathrm{E}^n f(\circ) - \mathrm{E} f(\circ) = \frac{1}{n} \sum_{k=1}^{n} [f(\circ, \xi_k(\omega)) - \mathrm{E} f(\circ)] \in K.$$

Очевидно, что $\mathrm{E}^n f(\circ) - \mathrm{E} f(\circ)$ можно рассматривать как случайные элементы, заданные на вероятностном пространстве (Ω, F, P) и принимающие значения в K.

В дальнейшем понадобятся следующие результаты из функционального анализа.

Определение 1 [14]. Пусть $(V, ||\circ||)$ — линейное нормированное пространство, $B(x, \rho)$ — замкнутый шар в V радиуса ρ с центром x, $f: V \to [-\infty, +\infty]$ — некоторая функция, $f(x_f) = \min\{f(x), x \in V\}$. Улучшающей функцией ψ для f в точке x_f называется монотонно неубывающая функция $\psi: [0, +\infty) \to \to [0, +\infty]$ с $\psi(0) = 0$ такая, что существует $\rho > 0$, для которого при всех $x \in B(x_f, \rho)$ имеем $f(x) \ge f(x_f) + \psi(||x - x_f||)$.

Пусть $V_0 \subset V$. Обозначим δ_{V_0} индикаторную функцию V_0:

$$\delta_{V_0}(x) = 0, \ x \in V_0, \ \delta_{V_0}(x) = +\infty, \ x \notin V_0.$$

Теорема 1 [14]. Пусть $(V, ||\circ||)$ — линейное нормированное пространство, $V_0 \subset V$ замкнуто; $f_0, g_0 : V \to \mathbb{R}$ — непрерывные функции на V. Обозначим

$$\varepsilon = \sup \{| f_0(x) - g_0(x)|,\ x \in V_0\}.$$

Введем функции $f, g : V \to (-\infty, +\infty]$ следующим образом: $f = f_0 + \delta_{V_0}$, $g = g_0 + \delta_{V_0}$. Тогда

$$| \inf \{f(x),\ x \in V\} - \inf \{g(x),\ x \in V\}| \le \varepsilon.$$

Далее, пусть x_f — точка минимума f: $f(x_f) = \inf \{f(x), x \in V\}$, ψ — улучшающая функция для f в точке x_f с некоторым радиусом $\rho > 0$. Если ε достаточно мало, так что для всех x, при которых $\psi(||x - x_f||) \le 2\varepsilon$, имеем $||x - x_f|| \le \rho$, то для любого $x_g \in \arg\min \{g(x), x \in B(x_f, \rho)\}$ справедливо неравенство $\psi(||x_f - x_g||) \le 2\varepsilon$.

Когда ψ — выпуклая и строго возрастающая на $[0, \rho]$ функция, предыдущее неравенство можно также выразить следующим образом: если ε настолько мало, что $\psi^{-1}(2\varepsilon) \le \rho$, то для любого $x_g \in \arg\min \{g(x), x \in B(x_f, \rho)\}$ получим

$$||x_f - x_g|| \le \psi^{-1}(2\varepsilon).$$

Имеет место следующее утверждение.

Теорема 2 [15]. Пусть $\{\mu_\varepsilon : \varepsilon > 0\}$ — семейство вероятностных мер на G, где G — замкнутое выпуклое подмножество сепарабельного банахова пространства E. Предположим, что предел $\Lambda(\lambda) = \lim\limits_{\varepsilon \to 0} \varepsilon \Lambda_{\mu_\varepsilon}(\lambda / \varepsilon)$ существует для любого $\lambda \in E^*$, где E^* — сопряженное к E пространство,

$$\Lambda_\mu(\lambda) = \ln\left(\int_E \exp[\langle \lambda, x \rangle] \mu(dx) \right)$$

для произвольной вероятностной меры μ на E, $\langle \lambda, x \rangle$ — соотношение двойственности.

Обозначим

$$\Lambda^*(q) = \sup \{\langle \lambda, q \rangle - \Lambda(\lambda), \lambda \in E^*\},\ q \in G.$$

Тогда функция Λ^* неотрицательна, полунепрерывна снизу, выпукла и для любого компактного множества $A \subset G$ выполнено

$$\limsup_{\varepsilon \to 0} \varepsilon \ln(\mu_\varepsilon(A)) \le -\inf \{\Lambda^*(q), q \in A\}.$$

Определение 2 [15]. Пусть Σ — сепарабельное банахово пространство, $\{\xi_i,\ i \in \mathbb{N}\}$ — стационарная в узком смысле случайная последовательность, определенная на вероятностном пространстве (Ω, F, P) и принимающая значения из Σ. Обозначим B_{mk} σ-алгебру подмножеств Ω, порожденную случайными элементами $\{\xi_i, m \le i \le k\}$. Для данного $l \in \mathbb{N}$ действительные случайные величины $\eta_1, \ldots, \eta_p$, $p \ge 2$, называются l-измеримо отделенными, если

$$-\infty \le m_1 \le k_1 < m_2 \le k_2 < \ldots < m_p \le k_p \le +\infty,$$

$$m_j - k_{j-1} \ge l,\ j = 2, \ldots, p,$$

и при каждом $j \in \{1, \ldots, p\}$ случайная величина η_j является $B_{m_j k_j}$-измеримой.

Определение 3 [15]. Случайная последовательность $\{\xi_i\}$ из определения 2 называется последовательностью с гиперперемешиванием, если существуют число $l_0 \in \mathbb{N} \cup \{0\}$ и невозрастающие функции $\alpha, \beta : \{l > l_0\} \to [1, +\infty)$ и $\gamma : \{l > l_0\} \to$ $\to [0, 1]$, удовлетворяющие соотношениям

$$\lim_{l \to \infty} \alpha(l) = 1,\ \limsup_{l \to \infty} l(\beta(l) - 1) < \infty,\ \lim_{l \to \infty} \gamma(l) = 0,$$

и для которых

$$||\eta_1 \dots \eta_p||_{L^1(P)} \le \prod_{j=1}^{p} ||\eta_j||_{L^{\alpha(l)}(P)} \tag{H-1}$$

при любых $p \ge 2$, $l > l_0$, $\eta_1, \dots, \eta_p$, l-измеримо отделенных, где

$$||\eta||_{L^r(P)} = \left(\int_{\Omega} |\eta(\omega)|^r \, dP \right)^{1/r},$$

$$\left| \int_{\Omega} \left(\xi(\omega) - \int_{\Omega} \xi(\omega) dP \right) \eta(\omega) dP \right| \le \gamma(l) ||\xi||_{L^{\beta(l)}(P)} ||\eta||_{L^{\beta(l)}(P)} \tag{H-2}$$

для всех $l > l_0$, $\xi, \eta \in L^1(P)$, l-измеримо отделенных.

Пусть X — компактное подмножество в $\mathbb{R}$. Как известно [16], $C(X)^* = M(X)$ — совокупность ограниченных знаковых мер на X,

$$\langle g, Q \rangle = \int_X g(x) Q(dx)$$

для любых $g \in C(X)$, $Q \in M(X)$.

Используя теорему 2, получаем следующее утверждение.

Теорема 3 [17]**.** Пусть $\{\xi_i,\ i \in \mathbb{N}\}$ — стационарная в узком смысле случайная последовательность, удовлетворяющая гипотезам (H-1) и (H-2) гиперперемешивания, или стационарная в узком смысле эргодическая случайная последовательность, удовлетворяющая гипотезе (H-1), определенная на вероятностном пространстве (Ω, F, P) и принимающая значения в компактном выпуклом множестве $K \subset C(X)$. Тогда для каждой меры $Q \in M(X)$ существует

$$\Lambda(Q) = \lim_{n \to \infty} \frac{1}{n} \ln \left(\int_{\Omega} \exp \left\{ \sum_{i=1}^{n} \int_X \xi_i(\omega)(x) Q(dx) \right\} dP \right)$$

и для любого замкнутого $A \subset K$

$$\limsup_{n \to \infty} \frac{1}{n} \ln \left(P \left\{ \frac{1}{n} \sum_{i=1}^{n} \xi_i \in A \right\} \right) \le -\inf \{\Lambda^*(g),\ g \in A\},$$

где $\Lambda^*(g) = \sup \left\{ \int_X g(x) Q(dx) - \Lambda(Q),\ Q \in M(X) \right\}$ — неотрицательная полунепрерывная снизу выпуклая функция.

В случае, когда $\{x_n, n \in \mathbb{N}\}$ принимает значения из $\mathbb{R}^d$, имеет место следующий результат, в котором вместо условия гиперперемешивания предполагается выполненным условие сильного перемешивания, которое приводится ниже.

Определение 4. Пусть $\{x_n,\ n \in \mathbb{N}\}$ — стационарная в узком смысле случайная последовательность. Определим σ-поле $\mathfrak{F}_{a,b} = \sigma\{x_k, a \le k \le b\}$ и пусть

$$\phi(n) = \sup \{|P(B / A) - P(B)| : A \in \mathfrak{F}_{-\infty, -n+1},\ B \in \mathfrak{F}_{1,\infty},\ P(A) > 0\}.$$

Полагаем, что $\{x_n, n \in \mathbb{N}\}$ удовлетворяет условию ϕ-перемешивания, если $\phi(n) \to 0$ при $n \to \infty$. Предполагаем также, что $e^{Kn} \phi(n) \to 0$ при $n \to 0$ для любого $K \ge 0$.

Теорема 4 [18]**.** Предположим, что $\{x_n, n \in \mathbb{N}\}$ является стационарной в узком смысле случайной последовательностью со значениями в пространстве $\mathbb{R}^d$, удовлетворяющей условию сильного перемешивания с $||x_k|| \le C < \infty$, $1 \le k \le n$. Пусть

$Z_n = (X_1 + \ldots + X_n) / n$, $n \ge 1$. Тогда существует полунепрерывная снизу функция $I : \mathbb{R}^d \to [0, \infty]$ с компактным множеством $I^{-1}[0, a]$, $a \ge 0$, и такая, что для каждого замкнутого множества $A \in \mathbb{R}^d$ выполнено

$$\limsup_{n \to \infty} n^{-1} \log P(Z_n \in A) \le -\inf_{x \in A} I(x),$$

а для любого открытого множества $A \in \mathbb{R}^d$ справедливо

$$\liminf_{n \to \infty} n^{-1} \log P(Z_n \in A) \le -\inf_{x \in A} I(x).$$

Кроме того, для каждого $\lambda \in \mathbb{R}^d$ существует предел

$$\lim_{n \to \infty} n^{-1} \log \mathrm{E}\{ \exp(n\lambda(Z_n)) \} = L(\lambda)$$

и функция $I(x)$ имеет вид

$$I(x) = \sup \{ \lambda(x) - L(\lambda) : \lambda \in \mathbb{R}^d \},$$

где $\lambda(x) = \sum \lambda_i x_i$.

Отметим также, что случай, когда $\{\xi_i, i \in \mathbb{N}\}$ в (2) является марковским процессом, рассматривался в работе [19], в которой получено утверждение, аналогичное теоремам 3 и 4.

Если имеют место утверждения теорем 3 или 4, то будем считать, что для последовательности $\{Z_n\}$ выполняется принцип больших уклонений. Теоремы 3 и 4 основополагающие при доказательстве утверждений о больших уклонениях для метода эмпирических средних теории стохастического программирования. В частности, используя теорему 3, получаем следующие утверждения.

Теорема 5 [17]**.** Пусть последовательность $\{\xi_i, i \in \mathbb{N}\}$ удовлетворяет условиям теоремы 3. Тогда для любого $\varepsilon > 0$ выполнено

$$\limsup_{n \to \infty} \frac{1}{n} \ln P\{ || F_n - \mathrm{E} f || \ge \varepsilon \} \le -\inf \{ I(z), z \in A_\varepsilon \},$$

где $A_\varepsilon = \{ z \in K : || z || \ge \varepsilon \}$;

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

— неотрицательная полунепрерывная снизу выпуклая функция;

$$\Lambda(Q) = \lim_{n \to \infty} \frac{1}{n} \ln \left(\int_\Omega \exp \left\{ \sum_{i=1}^n \int_X [f(x, \xi_i(\omega)) - \mathrm{E} f(x)] Q(dx) \right\} dP \right).$$

Теорема 6 [17]**.** Пусть выполнены условия теоремы 5. Тогда

$$\limsup_{n \to \infty} \frac{1}{n} \ln P\{ | \min \{ \mathrm{E} f(x), x \in X \} - \min \{ \mathrm{E}^n f(x), x \in X \} | \ge \varepsilon \} \le$$
$$\le -\inf \{ I(z), z \in A_\varepsilon \}, \tag{3}$$

где $I(\circ)$ и A_ε определены в теореме 5. Предположим, что существует улучшающая функция ψ для $\mathrm{E}f$ в точке x^* с некоторой постоянной ρ. Пусть x_n — точка минимума F_n по множеству $B(x^*, \rho)$. Если ε достаточно мало, так что выполнено условие

$$\psi(| x - x^* |) \le 2\varepsilon \Rightarrow | x - x^* | \le \rho,$$

где $\Rightarrow$ — знак логического следования, то

$$\limsup_{n \to \infty} \frac{1}{n} \ln P\{ \psi(| x_n - x^* |) \ge 2\varepsilon \} \le -\inf \{ I(z), z \in A_\varepsilon \}. \tag{4}$$

Доказательство. Согласно теореме 1 при каждом ω

$$|\min\{\mathrm{E}f(x), x\in X\}-\min\{\mathrm{E}^n f(x), x\in X\}|\le \|\mathrm{E}^n f-\mathrm{E}f\|.$$

Тогда из теоремы 5 следует справедливость неравенства (3). Далее, на основании теоремы 1 для всех ω выполнено $\psi(|x^*-x_n|)\le 2\|\mathrm{E}^n f-\mathrm{E}f\|$ и из теоремы 4 вытекает соотношение (4).

Следствие. Если, кроме условий теоремы 6, ψ выпукла и строго возрастает на $[0,\rho]$, то

$$\limsup_{n\to\infty}\frac{1}{n}\ln P\{|x_n-x^*|\ge\psi^{-1}(2\varepsilon)\}\le-\inf\{I(z), z\in A_\varepsilon\}. \tag{5}$$

Доказательство. В силу теоремы 1 для любого ω

$$|x_n-x^*|\le\psi^{-1}(2\|\mathrm{E}^n f-\mathrm{E}f\|).$$

Тогда

$$P\{|x_n-x^*|\ge\psi^{-1}(2\varepsilon)\}\le P\{\psi^{-1}(2\|\mathrm{E}^n f-Ef\|)\ge\psi^{-1}(2\varepsilon)\}=P\{\|\mathrm{E}^n f-\mathrm{E}f\|\ge\varepsilon\}$$

и справедливость (5) следует из теоремы 6.

Аналогично, с использованием теоремы 4, доказывается теорема о больших уклонениях для случая $Z_n\in\mathbb{R}^d$.

Далее рассмотрим случай непрерывной модели наблюдения. Объектом наблюдения является стационарный случайный процесс. Рассмотрим следующую задачу стохастической оптимизации: минимизировать функцию

$$F(x)=\mathrm{E}f(x,\xi(0)),\ x\in X, \tag{6}$$

где $\{\xi(t), t\in\mathbb{R}\}$ — стационарный в узком смысле эргодический случайный процесс с непрерывными траекториями, определенный на полном вероятностном пространстве (Ω, G, P), со значениями в некотором метрическом пространстве (Y,ρ), X — непустое компактное подмножество $\mathbb{R}$, $f: X\times Y\to\mathbb{R}$ — непрерывная функция.

Аппроксимируем задачу (6) следующей: минимизировать функцию

$$F_T(x)=\frac{1}{T}\int_0^T f(x,\xi(t))dt, \tag{7}$$

где $\{\xi(t), 0\le t\le T\}$ — наблюдения процесса $\xi(t)$, $T>0$.

Очевидно, что существует точка минимума $x_T(\omega)$ функции $F_T(x)$, являющаяся измеримой функцией.

Предположим, что $\mathrm{E}\max\{|f(x,\xi(0))|, x\in X\}<\infty$. Тогда функция (6) является непрерывной и у нее существует хотя бы одна точка минимума x_0. Предположим, что точка минимума единственна.

При выполнении приведенных выше условий с вероятностью 1 имеет место [13]

$$x_T\to x_0,\ F_T(x_T)\to F(x_0),\ T\to\infty.$$

Исследуем зависимость вероятности больших уклонений для x_T и минимального значения F_T соответственно от x_0 и $F(x_0)$.

Для любого y будем предполагать, что $f(\circ, y)$ является элементом пространства непрерывных функций $C(X)$. Предположим, что существует такое выпуклое компактное множество $K\subset C(X)$, что $f(\circ,y)-F(\circ)\in K$, $y\in Y$. Тогда $F_T(\circ)-F(\circ)\in K$.

Будем рассматривать F_T-F как случайные элементы на (Ω, G, P) со значениями в множестве K.

Понадобится следующее вспомогательное утверждение.

Теорема 7 [20]. Пусть $\{\xi(t), t \in \mathbb{R}\}$ — стационарный в узком смысле эргодический случайный процесс с непрерывными траекториями, удовлетворяющий гипотезе (H-1) гиперперемешивания на (Ω, G, P), принимающий значения в компактном выпуклом множестве $K \subset C(X)$. Тогда для каждой меры $Q \in M(X)$ существует

$$\Lambda(Q) = \lim_{T\to\infty} \frac{1}{T} \ln\left(\mathrm{E}\exp\left\{ \int_X \int_0^T \xi(t)(x)dtQ(dx) \right\} \right)$$

и для любого замкнутого $A \subset K$

$$\overline{\lim}\left\{ \frac{1}{T} \ln P\left\{ \frac{1}{T} \int_0^T \xi(t)dt \in A \right\}, T \to \infty \right\} \le -\inf\,\{\Lambda^*(g), g \in A\},$$

где $\Lambda^*(g) = \sup\left\{ \int_X g(x)Q(dx) - \Lambda(Q), Q \in M(X) \right\}$ — неотрицательная выпуклая полунепрерывная снизу функция.

Как и в случае дискретного времени, это утверждение является основным при доказательстве принципа больших уклонений для метода эмпирических средних при непрерывном параметре наблюдений. Имеет место следующее утверждение.

Теорема 8 [20]. Пусть процесс $\xi(t)$ удовлетворяет гипотезе (H-1) гиперперемешивания. Тогда при любом $\varepsilon > 0$

$$\overline{\lim}\left(\frac{1}{T} \ln P\{||F_T - F|| \ge \varepsilon\}, T \to \infty \right) \le -\inf\,\{I(z),\, z \in A_\varepsilon\},$$

где $A_\varepsilon = \{z \in K : ||z|| \ge \varepsilon\}$, $I(z) = \Lambda^*(z) = \sup\left\{ \int_X z(x)Q(dx) - \Lambda(Q), Q \in M(X) \right\}$ — неотрицательная полунепрерывная снизу выпуклая функция,

$$\Lambda(Q) = \lim_{T\to\infty} \frac{1}{T} \ln \mathrm{E}\exp \int_X \int_0^T (f(x, \xi(t)) - F(x))dtQ(dx)\,.$$

Доказательство проводится аналогично случаю дискретного времени [17].

Теорема 9. Пусть выполнены условия теоремы 8. Тогда

$$\overline{\lim}\left(\frac{1}{T} \ln P\left\{ |\min_{x\in X} F(x) - \min_{x\in X} F_T(x)| \ge \varepsilon \right\}, T \to \infty \right) \le -\inf\,\{I(z),\, z \in A_\varepsilon\},$$

где $I(\circ)$, A_ε определены в теореме 8.

Предположим, что существует улучшающая функция ψ для $F(x)$, определенной (6), в точке $x^* = \arg\min F(x)$ с некоторой постоянной r. Пусть x_T — точка минимума (7) по множеству $B(x^*, r)$. Если ε достаточно мало, так что выполнено условие

$$\psi(|x - x_0|) \le 2\varepsilon \Rightarrow |x - x_0| \le r,$$

то имеем

$$\overline{\lim}\left(\frac{1}{T} \ln P\{\psi(|x_T - x_0|) \ge 2\varepsilon\},\ T \to \infty \right) \le -\inf\,\{I(z),\ z \in A_\varepsilon\}\,.$$

Доказательство проводится так же, как доказательство теоремы 6.

Аналогично случаю дискретного времени имеет место следующее утверждение.

Следствие. Если, кроме условий теоремы 9, ψ выпукла и строго возрастает на $[0, r]$, то получаем

$$\overline{\lim}\left(\frac{1}{T} \ln P\{|x_T - x_0| \ge \psi^{-1}(2\varepsilon)\},\ T \to \infty \right) \le -\inf\,\{I(z), z \in A_\varepsilon\}\,.$$

Далее остановимся на случае, когда функция f зависит и от временно́го параметра. Рассмотрим дискретное время наблюдений. Предположим, что $\{\xi_i,\, i \in \mathbb{N}\}$ — стационарная в узком смысле эргодическая случайная последовательность, заданная на вероятностном пространстве (Ω, G, P) и принимающая значения в некотором измеримом пространстве $(Y, \aleph)$; $X = [a; b] \subset \mathbb{R}$; функция $\{f(i, x, y) : Z \times X \times Y \to \mathbb{R}\}$ выпукла по второму аргументу и измерима по третьему.

Рассмотрим задачу оптимизации

$$\min_{x \in X} \left\{ F_n(x) = \frac{1}{n} \sum_{i=1}^{n} f(i, x, \xi_i) \right\}.$$

Предположим, что:

1) при всех $i \in \mathbb{N}$, $x \in X$ $\mathrm{E}|f(i, x, \xi_0)| < \infty$;

2) для любого $x \in X$ существует функция $F(x) = \lim\limits_{n \to \infty} \mathrm{E} F_n(x)$;

3) найдутся такие $x^* \in X$, $c > 0$, что $F(x) \ge F(x^*) + c|x - x^*|$ $\forall x \in X$.

Из предположения 3 следует, что существует единственное решение задачи поиска $\min\limits_{x \in X} F(x)$, и это решение — точка x^*.

Очевидно, что при любых n, ω функция $F_n(\cdot)$ выпукла и для каждого n выпукла функция $\mathrm{E}F_n(\cdot)$.

Для произвольной функции $g : \mathbb{R} \to \mathbb{R}$ обозначим

$$g'_+(x) = \lim_{\Delta \to +0} \frac{g(x + \Delta) - g(x)}{\Delta}, \tag{8}$$

$$g'_-(x) = \lim_{\Delta \to +0} \frac{g(x - \Delta) - g(x)}{\Delta}, \tag{9}$$

если эти пределы существуют.

Введем обозначение

$$g_n(x) = \mathrm{E}F_n(x), \;\; x \in X.$$

Так как из выпуклости функции следует существование для нее пределов (8), (9), получаем, что такие пределы существуют:

- при всех i, y для функции $f(i, \cdot, y)$;
- при каждом i для $\mathrm{E}\, f(i, \cdot, \xi_0)$;
- при любых n, ω для $F_n(\cdot)$;
- при каждом n для $g_n(\cdot)$.

Для функций $\{f(i, x, y) : Z \times X \times Y \to \mathbb{R}\}$ используется следующее условие равномерного сильного перемешивания. Пусть

$$\alpha(j) = \sup_{\{i = 1, 2, \ldots\}} \; \sup_{\{(A, B) : A \in \sigma_{-\infty}^{i}, B \in \sigma_{i+j}^{\infty}\}} |P(AB) - P(A)P(B)|,$$

где $\sigma_n^m = \sigma\{f(i, \vec{u}, \xi_i),\, n \le i \le m,\, \vec{u} \in I\}$.

Полагаем, что функция $f(i, \vec{u}, \xi_i)$ удовлетворяет условию равномерного сильного перемешивания, если $\alpha(j) \to 0$ при $j \to \infty$.

В работе [21] приведены условия, когда

$$P\{x_n \to x^*,\, n \to \infty\} = 1, \tag{10}$$

$$P\{\mathrm{E}^n f(x_n) \to F(x^*),\, n \to \infty\} = 1. \tag{11}$$

Справедливо следующее утверждение.

Теорема 10. Пусть выполнены соотношения (10), (11) и следующие предположения:

1) случайная последовательность $\{\xi_i\}$ удовлетворяет условию гиперперемешивания (H-1);

2) существует такое $L > 0$, что для всех $i \in Z$, $y \in Y$

$$|f'_+(i, x^*, y)| \le L, \quad |f'_-(i, x^*, y)| \le L.$$

Тогда

$$\limsup_{n\to\infty} \frac{1}{n} \ln (P\{A_n^c\}) \le -\inf_{g \in F} \Lambda^*(g),$$

где $\Lambda^*(g) = \sup \{gQ(X) - \Lambda(Q), Q \in M(X)\}$,

$$\Lambda(Q) = \lim_{n\to\infty} \frac{1}{n} \ln \left(\int_\Omega \exp \left\{ Q(X) \sum_{i=1}^n \min [h'_+(i, x^*, \xi_i), h'_-(i, x^*, \xi_i)] \right\} dP \right),$$

$$A_n = \{\omega : \arg\min_{x \in X} F_n(x) = \{x^*\}\}, \; A_n^c = \Omega \setminus A_n.$$

Из данной теоремы для описанной выше нестационарной модели наблюдений следуют утверждения, аналогичные теоремам 8, 9 и следствию из теоремы 9.

Таким образом, получены утверждения о больших уклонениях для метода эмпирических средних с зависимыми наблюдениями, удовлетворяющие условиям гиперперемешивания или сильного перемешивания.

МЕТОД ЭМПИРИЧЕСКИХ СРЕДНИХ С РАЗРЫВНОЙ ЦЕЛЕВОЙ ФУНКЦИЕЙ

В данном разделе исследуется сходимость и скорость сходимости эмпирических аппроксимаций задачи в случае, когда случайная целевая функция $f(x, \xi_0)$ может быть разрывной по первому аргументу. Важным источником разрывных целевых функций в стохастическом программировании являются функции вероятности, квантили и другие индикаторы риска [22]. Функция вероятности выражает зависимость вероятности того или иного события от управляемых или оптимизируемых параметров. Задачи оптимизации вероятностей изучались в работах [23–26] Отметим, что оптимизация квантильной целевой функции сводится к оптимизации линейной функции при вероятностном ограничении [25, 26].

Задачи оптимизации вероятностей, вообще говоря, являются невыпуклыми и поэтому вычислительно трудными. Условия выпуклости таких задач детально исследованы в [24–27]. В работах [28–37] задачи оптимизации вероятностей с дискретным распределением случайных данных сводились к детерминированным задачам смешанного целочисленного программирования, которые затем численно решались современными пакетами программ дискретной оптимизации. Для реализации такого подхода применительно к задачам с непрерывным распределением случайных данных необходимо аппроксимировать случайную вероятностную меру дискретной (например, эмпирической) мерой и обеспечить сходимость аппроксимаций функции вероятности к исходной функции. Аналогичный подход под названием метода эмпирических средних широко используется в стохастическом программировании для оптимизации функции среднего значения [5, 11–13]. Сложность в реализации подобного плана для оптимизации функций вероятности заключается в том, что последние в общем случае являются разрывными и поэтому нельзя гарантировать равномерной сходимости аппроксимаций и, следовательно, сделать заключение о сходимости минимумов приближенных задач к минимумам исходной задачи.

Эту трудность преодолевают следующим образом. Вместо равномерной сходимости устанавливается эпи-сходимость аппроксимаций [36], т.е. сходимость надграфиков последовательности аппроксимаций к надграфику исходной функции вероятности. Этого свойства достаточно для сходимости оптимумов. Доказательство основано на законе больших чисел для случайных полунепрерывных функций [38].

Пусть $(\Omega, \Sigma, \mathrm{P})$ — некоторое вероятностное пространство, $X: \Omega \to \mathrm{X}(\Omega) \subseteq \subseteq \mathbb{R}^m$ — векторная случайная величина со значениями в множестве $\mathrm{X}(\Omega)$, $U \subseteq \mathbb{R}^n$ — множество допустимых стратегий оптимизации, $\Phi{:}U \times \mathrm{X}(\Omega) \to \mathbb{R}^1$ и $Q{:}\ U \times \mathrm{X}(\Omega) \to \mathbb{R}^1$ — борелевские по $x \in \mathrm{X}(\Omega)$ функции для всех $u \in U$. Определим функции математического ожидания

$$f_1(u) \overset{\text{def}}{=} \mathrm{E}[\Phi(u, X)] = \int_\Omega \Phi(u, X(\omega)) d\mathrm{P}(\omega),$$

вероятности

$$P_\varphi(u) \overset{\text{def}}{=} \mathrm{P}\{\Phi(u, X) \le \varphi, Q(u, X) \le 0\} = \mathrm{P}\{\omega \in \Omega : \Phi(u, X(\omega)) \le \varphi, Q(u, X(\omega)) \le 0$$

и квантили

$$\varphi_\alpha(u) \overset{\text{def}}{=} \inf\{\varphi \in \mathbb{R}^1 : P_\varphi(u) \ge \alpha\} = \min\{\varphi \in \mathbb{R}^1 : \mathrm{P}\{\Phi(u, X) \le \varphi, Q(u, X) \le 0\} \ge \alpha\},$$

где α — параметр, $0 < \alpha < P^*$, $\mathrm{P}\{\cdot\}$ — вероятность события в скобках (по определению $\mathrm{P}\{\varnothing\} = 0$), E — знак математического ожидания,

$$P^* \overset{\text{def}}{=} \sup_{u \in U} [P(u) = \mathrm{P}\{\omega \in \Omega{:}\ Q(u, X(\omega)) \le 0\}]. \tag{12}$$

Известно, что задача минимизации функции квантили эквивалентна следующей задаче (см. [25, sec. 4.2]):

$$\varphi \to \inf_{\varphi \in \mathbb{R}^1,\, u \in U},$$

$$P_\varphi(u) = \mathrm{P}\{\Phi(u, X) - \varphi \le 0,\ Q(u, X) \le 0\} \ge \alpha.$$

Заметим, что функция $P_\varphi(\cdot)$ определена на всем множестве U. Свойства функций вероятности и квантили детально изучены [25, 26]. Так, если $\Phi(u, x)$, $Q(u, x)$ полунепрерывны снизу по u при каждом x, то функция $P_\varphi(u)$ полунепрерывна сверху по совокупности переменных (u, φ). Кроме того, функция $P_\varphi(u)$ монотонна (не убывает) по φ и непрерывна справа, поэтому инфимум в определении $\varphi_\alpha(u)$ достигается. Функция квантили является специальным случаем маргинальной функции (функции минимума). При сделанных предположениях она полунепрерывна снизу по (u, α).

Пусть $\{X(\omega) = X_1(\omega), X_2(\omega), \ldots\}$ — последовательность независимых одинаково распределенных случайных величин. Рассмотрим следующую последовательность аппроксимаций $P_N(u)$ функции $P(u)$:

$$P_N(u) = \frac{1}{N} \sum_{k=1}^{N} \chi(Q(u, X_k)), \tag{13}$$

где

$$\chi(Q(u, X_k)) = \begin{cases} 1, & Q(u, X_k) \le 0, \\ 0, & Q(u, X_k) > 0. \end{cases}$$

Здесь для краткости записи используем обозначение $P_N(u)$, хотя эта функция зависит также от $X_1, \ldots, X_N$. Будем предполагать, что функции $Q(u, x)$ полунепрерывны снизу по u при каждом x (и измеримы по Борелю по (u, x)), тогда функции $\chi(Q(u, X_k))$ и $P_N(u)$ полунепрерывны сверху по u. Наряду с (12) рассмотрим последовательность задач

$$P_N^* = \sup_{u \in U} \left[P_N(u) = \frac{1}{N} \sum_{k=1}^{N} \chi(Q(u, X_k)) \right]. \tag{14}$$

Необходимо доказать, что $\lim_{N\to\infty} P_N^* = P^*$ (с вероятностью единица) при том, что функции $P_N(u)$ разрывные, а также указать эффективный способ решения разрывных задач (14).

В работе [33] предложен следующий подход к решению задачи (аналогичный подход к задаче с вероятностными ограничениями описан в [28–30]). Предположим, что $Q(u,x)\le M_x$ для всех $u\in U$. Задача (14) эквивалентна следующей задаче смешанного целочисленного программирования:

$$\sum_{k=1}^{N}\delta_k \to \max_{\{u\in U,\ \delta_k\in\{0,1\}\}} \tag{15}$$

при ограничениях

$$Q(u,X_k)\le M_{X_k}(1-\delta_k),\ k=1,\dots,N, \tag{16}$$

где M_{X_k} выбраны так, чтобы $M_{X_k}\ge \sup_{u\in U} Q(u,X_k)$. При линейной функции $Q(\cdot,x)$ и полиэдральном множестве U задачу (15), (16) можно эффективно решить современными пакетами программ дискретной оптимизации [36]. Доказательства эквивалентности задач (14) и (15), (16), а также дальнейшие обобщения этого подхода изложены в [33–35, 37].

Перейдем к обоснованию метода эмпирических средних для оптимизации функций вероятности. Одна из возможностей состоит в установлении условий равномерной сходимости (в некотором вероятностном смысле) функций $P_N(u)$ из (14) к функции $P(u)$ из (12), тогда из равномерной сходимости будет следовать сходимость оптимальных значений и решений.

Пусть X, Z — метрические пространства; (Ω,Σ_Ω,P) — полное вероятностное пространство; $f:X\times Z\to\mathbb{R}^1$ — измеримая по Борелю функция двух переменных (x,z); $\{\xi,\xi_1,\xi_2,\dots\}$ — последовательность независимых одинаково распределенных величин со значениями в пространстве Z. Согласно равномерному закону больших чисел

$$\delta_n(\xi^n)=\sup_{x\in X}\left|\frac{1}{n}\sum_{i=1}^{n}f(x,\xi_i)-\mathrm{E}_P f(x,\xi)\right|\to 0,\ n\to\infty, \tag{17}$$

в некотором вероятностном смысле при условии измеримости величин $\delta_n(\xi^n)$. Используя технику симметризации (см., например, [39]), можно получить оценку

$$\mathrm{E}_P\delta_n(\xi^n)\le 2\mathrm{E}_P\mathrm{E}_\sigma\max_{x\in X}\left|\frac{1}{n}\sum_{i=1}^{n}\sigma_i f(x,\xi_i)\right|, \tag{18}$$

где $\{\sigma_1,\dots,\sigma_n\}$ — независимые (радемахеровские) случайные величины, принимающие значения ± 1 с вероятностями 0.5, E_σ — математическое ожидание по $\{\sigma_1,\dots,\sigma_n\}$. Величина

$$R_n(f,z^n)=\mathrm{E}_\sigma\sup_{x\in X}\left|\frac{1}{n}\sum_{i=1}^{n}\sigma_i f(x,z_i)\right|$$

называется средним Радемахера для множества $A(z^n)=\{a=(f(x,z_1),\dots$ $\dots,f(x,z_n))\in\mathbb{R}^n, x\in X\}$, где $z^n=(z_1,\dots,z_n)$. При условии $\lim\limits_{n\to\infty}\mathrm{E}R_n(f,\xi^n)=0$ из неравенства (18) следует равномерный закон больших чисел (17) (в среднем, а также по вероятности). Базовые свойства и исчисление радемахеровских средних для множеств рассмотрены в [39]. В работах [40, 41] эти результаты переформулированы и обобщены на семейства ограниченных функций $\{f(\cdot,z)\}_{z\in Z}$, определенных на X. В случае эмпирической аппроксимации (13) функции вероятности $P_N(u)$ из (12) обозначим $f(x,z)=\chi(Q(x,z))$ и рассмотрим радемахеровские средние

$$R_n(z^n)=\mathrm{E}_\sigma\sup_{x\in X}\left|\frac{1}{n}\sum_{i=1}^{n}\sigma_i\chi(Q(x,z_i))\right|.$$

Оценки для $R_n(z^n)$ из [41] для разрывной функции $\chi(Q(\cdot, z))$ с учетом того, что $|\chi(Q(\cdot, z))| \le 1$, принимают следующий вид.

Лемма 1 (оценки радемахеровских средних)**.** Для дискретного множества X с числом элементов $|X|$ имеет место оценка $R_n(z^n) \le \sqrt{2\ln|X|/n}$. Для дискретного множества Z с числом элементов $|Z|$ имеет место оценка $R_n(z^n) \le \sqrt{|Z|/n}$. Если $Q(x, z)$ квазивогнута или квазивыпукла (в частности, монотонна) по $z \in Z \subset \mathbb{R}^1$ для любого $x \in X$, то $R_n(z^n) \le 2\sqrt{\ln n/n}$.

Оба рассмотренных в лемме 1 случая (дискретность X или Z) представляют практический интерес. Дискретное множество X возникает в широко распространенных задачах дискретного стохастического программирования, а дискретное множество Z появляется, когда многократная независимая выборка $\{\xi_i \in Z, i = 1, \ldots, n\}$ осуществляется из фиксированного набора исторических наблюдений.

Оценки леммы 1 гарантируют в силу (18) выполнение равномерного закона больших чисел в среднем и по вероятности. Кроме того, лемма 1 дает оценки скорости равномерной сходимости. При конечном множестве X равномерный закон больших чисел, очевидно, выполняется и с вероятностью единица.

Другие условия выполнения равномерного функционального закона больших чисел формулируются в терминах сложности Вапника–Червоненкиса для класса функций $\{f(\cdot, z)\}_{z \in Z}$ [40, 42].

Качественное обоснование сходимости метода эмпирических средних в разрывном случае возможно на основе усиленного закона больших чисел для случайных полунепрерывных снизу функций [36, 38]. Далее приведем качественное обоснование данного метода на основе так называемого графического закона больших чисел для случайных компактнозначных отображений [43, 44].

Рассмотрим также следующие многозначные отображения:

$$\bar{P}(u) = \mathrm{E}S(u, \omega), \quad S(u, \omega) = \begin{cases} 1, & Q(u, X) < 0, \\ [0,1], & Q(u, X) = 0, \\ 0, & Q(u, X) > 0, \end{cases}$$

$$\bar{P}_N(u) = \frac{1}{N}\sum_{k=1}^{N} S_k(u, \omega), \quad S_k(u, \omega) = \begin{cases} 1, & Q(u, X_k) < 0, \\ [0,1], & Q(u, X_k) = 0, \\ 0, & Q(u, X_k) > 0, \end{cases}$$

где $\mathrm{E}S(u, \omega)$ обозначает (при каждом фиксированном u) математическое ожидание случайного множества $S(u, \omega)$ в смысле [45], т.е. совокупность интегралов от однозначных измеримых селекторов отображения $S(u, \omega)$, а $\bar{P}_N(u)$ является (поэлементной) суммой множеств $S_k(u, \omega)$.

Рассмотрим следующие задачи оптимизации многозначных функций $\bar{P}(u)$ и $\bar{P}_N(u)$:

$$\bar{P}^* = \sup\{p \in \bar{P}(u) : u \in U\}, \tag{19}$$

$$\bar{P}_N^* = \sup\{p \in \bar{P}_N(u) : u \in U\}. \tag{20}$$

Задачи (19), (20) можно трактовать как задачи оптимизации линейной функции на графиках $\mathrm{gph}\,\bar{P}$, $\mathrm{gph}\,\bar{P}_N$ многозначных отображений $\bar{P}(\cdot)$ и $\bar{P}_N(\cdot)$. Заметим, что в сделанных предположениях задачи (12) и (19), а также (14) и (20) имеют решения и эквивалентны в следующем смысле.

Определение 5 [34]**.** Две задачи оптимизации эквивалентны, если выполнены следующие условия:

а) либо обе задачи имеют допустимые решения (с конечными значениями целевых функций), либо обе не имеют таких решений;

б) если эти задачи имеют допустимые решения, то оптимальные значения их целевых функций (конечные или бесконечные) совпадают;

в) если оптимальные значения их целевых функций конечны, то эти значения в обеих задачах либо достигаются, либо не достигаются;

г) если оптимальные значения достигаются, то по оптимальному решению одной задачи с помощью явно указанного алгоритма за конечное число шагов восстанавливается оптимальное решение другой.

Доказательство эквивалентности оптимизационных задач основано на установлении специального соответствия между множествами допустимых стратегий задач, а именно, для каждой допустимой стратегии одной задачи указываем допустимую стратегию другой с таким же или лучшим значением целевой функции.

Лемма 2 [34]**.** Две оптимизационные задачи

$$\Phi_1(u) \to \sup_{u \in U_1}, \ \Phi_2(u) \to \sup_{u \in U_2}$$

эквивалентны в смысле определения 1, если известны алгоритмы (отображения) $A_1 : U_1 \to U_2$ и $A_2 : U_2 \to U_1$, которые для каждой допустимой стратегии одной задачи указывают допустимую стратегию другой задачи с таким же или бóльшим значением целевой функции, т.е. для любых $u_1 \in U_1$ и $u_2 \in U_2$ выполнено $\Phi_2(A_1(u_1)) \ge \Phi_1(u_1)$ и $\Phi_1(A_2(u_2)) \ge \Phi_2(u_2)$.

При этом, если u_1^* — оптимальное решение первой задачи, то $A_1(u_1^*)$ — оптимальное решение второй задачи и $\Phi_1(u_1^*) = \Phi_2(A_1(u_1^*))$. Наоборот, если u_2^* — оптимальное решение второй задачи, то $A_2(u_2^*)$ — оптимальное решение первой задачи и $\Phi_2(u_2^*) = \Phi_1(A_2(u_2^*))$.

Лемма 3. В сделанных предположениях задачи (12) и (19), а также задачи (14) и (20) имеют решения и эквивалентны.

Далее потребуется использование графического закона больших чисел для компактнозначных случайных многозначных отображений.

Теорема 11 [43, теорема 3.1]. Предположим, что:

(i) $\overline{U}$ — замкнутое множество в евклидовом пространстве $\mathbb{R}^n$, B_n — борелеская σ-алгебра в $\mathbb{R}^n$, (Ω, Σ, P) — полное вероятностное пространство;

(ii) отображения $S(u,\omega) = S_0(u,\omega)$, $S_k(u,\omega) : U \times \Omega \to 2^{\mathbb{R}^m}$, $k = 0,1,\ldots$, компактнозначны, $B_n \times \Sigma$-измеримы по (u,ω) и полунепрерывны сверху по $u \in \overline{U}$ для P-почти всех $\omega \in \Omega$, т.е. их графики $\operatorname{gph} S_k(\cdot,\omega)$ — замкнутые случайные множества в $\overline{U} \times \mathbb{R}^m$;

(iii) случайные множества (графики) $\{\operatorname{gph} S_k(\cdot,\omega)\}$ независимы и одинаково распределены как $\operatorname{gph} S(\cdot,\omega)$;

(iv) для любого компактного подмножества $U \subseteq \overline{U}$ существует интегрируемая функция $M_U : \Xi \to \mathbb{R}^1$ такая, что $|| S_i(u,\omega) || \le M_U(\omega) \ \forall i, \forall (u,\omega) \in U \times \Omega$.

Обозначим $S^N(u) = N^{-1} \sum_{k=1}^{N} S_k(u,\omega)$, где под суммой множеств понимается поэлементная сумма. Тогда отображение $\overline{S}(u) = \mathrm{E}\operatorname{co}\{S(\cdot,\omega)\}$ полунепрерывно сверху и графики отображений S^N сходятся к графику $\overline{S}$, т.е. для любого компактного подмножества $U \subseteq \overline{U}$ выполнено

$$\lim_{N \to \infty} D_H(\operatorname{gph} S^N(\cdot,\omega), \ \operatorname{gph} \mathrm{E}\operatorname{co}\{S(\cdot,\omega)\}) = 0 \ \ P\text{-п.н.},$$

где $D_H(\cdot,\cdot)$ — расстояние Хаусдорфа между множествами, Е — математическое ожидание (интеграл по мере P в смысле Аумана [45]), $\operatorname{co}\{\cdot\}$ — выпуклая оболочка.

Лемма 4 (сходимость графиков отображений). Для компактного множества U выполнено

$$\lim_{\nu \to \infty} D_H (\operatorname{gph} \overline{P}_N (\cdot, \omega), \operatorname{gph} \overline{P}(\cdot)) = 0 \ \ P\text{-п.н.}$$

Данная лемма — прямое следствие теоремы 11.

Лемма 5 (сходимость максимумов (20) к максимумам (19). Имеет место

$$\operatorname*{Lim}_{N \to \infty} D_H \left(\operatorname*{arg\,max}_{\{p \in \overline{P}_N(u),\, u \in U\}} \overline{P}_N (u), \ \operatorname*{arg\,max}_{\{p \in \overline{P}(u),\, u \in U\}} \overline{P}(u) \right) = 0$$

и, как следствие, выполнено $\lim_{N \to \infty} \overline{P}_N^* = \overline{P}^*$.

Утверждение леммы вытекает из сходимости компактных множеств $\operatorname{gph} \overline{P}_N (\cdot, \omega)$ к компактному множеству $\operatorname{gph} \overline{P}(\cdot)$.

В данном разделе обоснован метод эмпирических средних для оптимизации, вообще говоря, разрывных функций вероятности. Функция вероятности является одним из основных критериев при оптимизации стохастических систем. Метод состоит в аппроксимации функций вероятности их эмпирическими оценками, сведении приближенных задач к задачам смешанного целочисленного программирования и решении последних современными пакетами прикладных программ дискретной оптимизации. Проблема с обоснованием данного подхода к оптимизации функций вероятности заключается в том, что аппроксимации функций вероятности разрывные и в общем случае не сходятся равномерно к, возможно, разрывной исходной функции вероятности. Обоснование метода состоит в доказательстве сходимости расширенных графиков аппроксимирующих функций к расширенному графику исходной функции вероятностей. Сходимость графиков является следствием закона больших чисел для случайных многозначных отображений.

ЗАКЛЮЧЕНИЕ

Резюмируя представленные результаты, отметим, что их можно использовать при решении новых классов задач стохастической оптимизации, которые возникают в теории распознавания, регрессионного анализа и других областях, где появляется необходимость нахождения оптимальных решений при нетрадиционных условиях на функции цели и зависимость наблюдений.

СПИСОК ЛИТЕРАТУРЫ

1. Dantzig G.B., Thapa M.N. Linear programming 2: Theory and extensions. New York: Springer, 2003. Vol. 2. 448 p.
2. Numerical techniques for stochastic optimization. Yu. Ermoliev, R.J-B. Wets, Eds. Berlin: Springer, 1988. 571 p.
3. Birge J., Luveaux F. Introduction to stochastic programming. New York: Springer, 1997. 421 p.
4. Stochastic programming. A. Ruszczyński, A. Shapiro, Eds. Handbooks in OR&MS. Amsterdam: Elsevier, 2003. Vol. 10. 682 p.
5. Shapiro A., Dentcheva D., Ruszczyński A. Lectures on stochastic programming: Modeling and theory. Philadelphia: SIAM, 2009. 442 p.
6. Ермольев Ю.М. Модели и методы стохастического программирования. Москва: Наука, 1976. 240 с.
7. Юдин Д.Б. Задачи и методы стохастического программирования. Москва: Сов. радио, 1979. 392 с.
8. Ermoliev Yu. Stochastic quasigradient methods. Numerical techniques for stochastic optimization. Yu. Ermoliev, R. Wets, Eds. Berlin: Springer, 1988. P. 141–185.
9. Pflug G.Ch. Optimization of stochastic models. The interface between simulation and optimization. Boston; Dordrecht; London: Kluwer Academic Publishers, 1996. 382 p.
10. Kushner H.J., Yin G.G. Stochastic approximation and recursive algorithms and applications. Sec. ed. New York: Springer, 2003. 474 p.

11. Nemirovski A., Juditsky A., Lan G., Shapiro A. Robust stochastic approximation approach to stochastic optimization. *SIAM J. Optim*. 2005. Vol. 19, N 4. P. 1574–1609.

12. Shapiro A. Monte Carlo sampling methods. Stochastic programming. A. Ruszczyński, A. Shapiro, Eds. Handbooks in OR&MS. Amsterdam: Elsevier, 2003. Vol. 10. P. 353–425.

13. Knopov P.S., Kasitskaya E.J. Empirical estimates in stochastic optimization and identification. Dordrecht; Boston; London: Kluwer Academic Publishers, 2002. 250 p.

14. Kaniovski Yu.M., King A.J., Wets R.J.-B. Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems. *Ann. Oper. Res*. 1995. Vol. 56. P. 189–208.

15. Deuschel J.-D., Stroock D.W. Large deviations. Boston: Acad. Press, Inc., 1989. 310 p.

16. Dunford N., Schwartz J. Linear operators, part I: General theory. New York: Interscience, 1957. 896 p.

17. Кнопов П.С., Касицкая Е.И. О больших уклонениях эмпирических оценок в задачах стохастического программирования. *Кибернетика и системный анализ*. 2004. № 4. С. 52–61.

18. Bryc W. On large deviations for uniformly strong mixing sequences. *Stochastic Processes and their Applications*. 1992. Vol. 41. P. 191–202.

19. Dal L., Chen C.H., Birge J.R. Convergence properties of two-stage stochastic programming. *J. Optim. Theory Appl*. 2000. Vol. 106, N 3. P. 489–509.

20. Knopov P., Kasitskaya E. Large deviations for the method of empirical means in stochastic optimization problems with continuous time observations. Optimization Methods and Applications. In Honor of the 80-th Birthday of Ivan V. Sergienko. S. Butenko, P.M. Pardalos, and V. Shylo, Eds. New York: Springer, 2017. P. 263–275.

21. Кнопов П.С., Касицкая Е.И. О больших уклонениях эмпирических оценок в задаче стохастического программирования при нестационарных наблюдениях. *Кибернетика и системный анализ*. 2010. Т. 46, № 5. С. 46–50.

22. Ermoliev Y., Norkin V. Stochastic optimization of risk functions via parametric smoothing. *Dynamic Stochastic Optimization*. K. Marti, Y. Ermoliev, G. Pflug, Eds. Berlin; Heidelberg; New York: Springer-Verlag, 2004. P. 225–247.

23. Райк Э. Качественные исследования в задачах стохастического нелинейного программирования. *Изв. АН ЭССР. Сер. физ.-мат*. 1971. Т. 20, № 1. С. 8–14.

24. Prékopa A. Stochastic programming. Dordrecht: Kluwer Academic Publishers, 1995. 600 p.

25. Kibzun A.I., Kan Y.S. Stochastic programming problems with probability and quantile functions. Chichester; New York; Brisbane: John Wiley&Sons, 1996. 301 p.

26. Кан Ю.С., Кибзун А.И. Задачи стохастического программирования с вероятностными критериями. Москва: Физматлит, 2009. 372 с.

27. Норкин В.И., Роенко Н.В. α-вогнутые функции и меры и их применения. *Кибернетика и системный анализ*. 1991. № 6. С. 77–88.

28. Корбут А.А., Финкельштейн Ю.Ю. Дискретное программирование. Москва: Наука, 1969. 368 с.

29. Sen S. Relaxation for probabilistically constrained programs with discrete random variables. *Operations Research Letters*. 1992. Vol. 11. P. 81–86.

30. Ruszczyński A. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedral. *Math. Program*. 2002. Vol. 93. P. 195–215.

31. Benati S., Rizzi R. A mixed integer linear programming formulation of the optimal mean–Value-at-Risk portfolio problem. *Eur. J. Oper. Res*. 2007. Vol. 176. P. 423–434.

32. Luedtke J., Ahmed S., Nemhauser G. An integer programming approach for linear programs with probabilistic constraints. *Math. Program*. 2010. Vol. 122, N 2. P. 247–272.

33. Норкин В.И., Бойко С.В. Оптимизация финансового портфеля на основе принципа безопасности. *Кибернетика и системный анализ*. 2012. Т. 48, № 2. С. 29–41.

34. Кибзун А.И., Наумов А.В., Норкин В.И. О сведении задачи квантильной оптимизации с дискретным распределением к задаче смешанного целочисленного программирования. *Автоматика и телемеханика*. 2013. № 6. С. 66–86.

35. Норкин В.И., Кибзун А.И., Наумов А.В. Сведение задач двухэтапной вероятностной оптимизации с дискретным распределением случайных данных к задачам частично целочисленного программирования. *Кибернетика и системный анализ*. 2014. Т. 50, № 5. С. 34–48.

36. Pagnoncelli B.K., Ahmed S., Shapiro A. Sample average approximation method for chance constrained programming: Theory and applications. *J. Optim. Theory Appl*. 2009. Vol. 142. P. 399–416.

37. Luedtke J. A branch-and-cut decomposition algorithm for solving chance-constrained mathematical programs with finite support. *Mathematical Programming*. 2014. Vol. 146, Iss. 1–2. P. 219–244.

38. Artstein Z., Wets R.J.-B. Consistency of minimizers and the SLLN for stochastic programs. *J. Convex Anal*. 1996. Vol. 2. P. 1–17.

39. Boucheron S., Bousquet O., Lugosi G. Theory of classification: A survey of some recent advances. *ESAIM: Probability and Statistics*. 2005. Vol. 9. P. 323–375.

40. Pflug G.Ch. Stochastic optimization and statistical inference. Stochastic programming, A. Ruszczyński, A. Shapiro, Eds. Handbooks in OR&MS. Amsterdam: Elsevier, 2003. Vol. 10. P. 427–482.

41. Ermoliev Y.M., Norkin V.I. Sample average approximation method for compound stochastic optimization problems. *SIAM J. Optim*. 2013. Vol. 23, N 4. P. 2231–2263.

42. Vapnik V.N. Statistical learning theory. New York: Wiley, 1998. 736 p.

43. Norkin V.I., Wets R.J-B. On a strong graphical law of large numbers for random semicontinuous mappings. *Вестник СПбГУ. Сер. 10*. 2013. Вып. 3. С. 102–111.

44. Shapiro A., Xu H. Uniform law of large numbers for set-valued mappings and subdifferentials of random functions. *J. Math. Anal. Appl*. 2007. Vol. 325. P. 1390–1399.

45. Aumann R.J. Integrals of set-valued functions. *J. Math. Anal. Appl*. 1965. Vol. 12, N 1. P. 1–12.

Надійшла до редакції 06.07.2017

П.С. Кнопов, В.І. Норкін

ПРО УМОВИ ЗБІЖНОСТІ МЕТОДУ ЕМПІРИЧНИХ СЕРЕДНІХ У СТОХАСТИЧНОМУ ПРОГРАМУВАННІ

Анотація. Проаналізовано умови збіжності методу емпіричних середніх у стохастичному програмуванні за нетрадиційних умов, коли використовуються залежні спостереження випадкових параметрів задачі та випадкові показники оптимізації можуть бути розривними індикаторними функціями. Для випадку залежних спостережень встановлено теореми про ймовірності великих відхилень для наближених оптимальних значень та розв'язків.

Ключові слова: стохастичне програмування, метод емпіричних середніх, залежні спостереження, умови перемішування, великі відхилення, розривні функції, функції ймовірності, збіжність методу.

P.S. Knopov, V.I. Norkin

ABOUT CONVERGENCE CONDITIONS FOR THE EMPIRICAL MEAN METHOD OF STOCHASTIC PROGRAMMING

Abstract. The paper analyzes convergence conditions of the empirical mean method under nonstandard conditions, where dependent observations of random parameters are used and probabilistic optimization functions may be discontinuous indicators. For the case of dependent observations, large deviation type theorems for approximate optimal values and solutions are established.

Keywords: stochastic programming, empirical mean method, mixing conditions, large deviations, discontinuous functions, probability functions, method convergence.

Кнопов Павел Соломонович,
чл.-кор. НАН Украины, профессор, заведующий отделом Института кибернетики им. В.М. Глушкова НАН Украины, Киев, e-mail: knopov1@yahoo.com.

Норкин Владимир Иванович,
доктор физ.-мат. наук, ведущий научный сотрудник Института кибернетики им. В.М. Глушкова НАН Украины; профессор кафедры Национального технического университета Украины «Киевский политехнический институт имени Игоря Сикорского», e-mail: vladimir.norkin@gmail.com.