УДК 519.853.4

В.И. НОРКИН, Б.О. ОНИЩЕНКО

О СТОХАСТИЧЕСКОМ АНАЛОГЕ МЕТОДА ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ ПИЯВСКОГО

В работе обоснован метод Пиявского для решения задачи стохастической глобальной оптимизации с целевой функцией типа математического ожидания. В частности, показано, что в качестве касательных минорант для таких функций можно брать математическое ожидание стохастических касательных минорант подынтегральной функции. Обоснованы модификации метода, связанные с использованием не касающихся минорант.

Введение. Метод Пиявского [1–3] неоднократно переоткрывался и является одним из популярных методов глобальной оптимизации [4]. Он имеет две эквивалентные формы: для оптимизации функций максимума и функций, допускающих так называемые касательные миноранты [5]. Понятие касательных минорант является ключевым для данного метода. В случае липшицевых функций касательные миноранты имеют вид касательных конусов. Однако существует много других видов касательных минорант, например, касательные параболоиды и др. [5]. В настоящей работе метод распространяется на задачи стохастической глобальной оптимизации, рассматриваются некоторые модификации исходного метода Пиявского и производится численное исследование эффективности метода при использовании касательных конусов и параболоидов.

Стохастические касательные миноранты. Рассмотрим задачу стохастической глобальной оптимизации:

$$\min_{x\in X}\,[F(x)=Ef(x,\theta)], \tag{1}$$

где θ – случайный параметр; E – символ математического ожидания по θ, $f(x,\theta)$ – некоторая непрерывная по x и интегрируемая по θ функция; $\theta\in\Theta$; X – непрерывное или дискретное множество.

Определение 1 [5]. Пусть X – топологическое пространство, функции $F(x), x\in X$, и $\varphi(x,y), x\in X, y\in X$, связаны условиями:

(i) $F(x)\ge\varphi(x,y)$ для всех $x\in X$, $y\in X$;

(ii) $F(y)=\varphi(y,y)$ для всех $y\in X$;

(iii) функция $\varphi(x,y)$ непрерывна по x равностепенно по y (см. [5]).

Тогда функции $\{\varphi(\cdot,y),\ y\in X\}$ называются касательными (в точках y) минорантами для $F(x)$.

Лемма 1. Если функция $\varphi(x,y)$ непрерывна по совокупности переменных (x,y), то она равномерно непрерывна по (x,y) и, следовательно, непрерывна по x равностепенно по y.

Таким образом, если $\lim_{k\to\infty}\varphi(x^k,y^k)=\varphi(x,y)$ для любых $x^k\to x$, $y^k\to y$, то функция $\varphi(x,y)$ непрерывна по x равностепенно по y.

Примечание. Исчисление касательных минорант построено в [5]. В работе [6] рассматриваются вогнутые, возможно, разрывные по x миноранты.

Определение 2. Функции $\{\varphi(\cdot,y,\theta),\ y\in X,\theta\in\Theta\}$, где Θ – носитель некоторого вероятностного пространства (Θ,Σ,P), называются стохастическими касательными минорантами для $F(x)$, если функции $\varphi(x,y,\theta)$ измеримы по θ, а математические ожидания $\varphi(x,y)=E\varphi(x,y,\theta)$ конечны и для каждого $y\in X$ являются касательными в точке y минорантами для $F(x)$.

Лемма 2. Предположим, что функции $f(\cdot,\theta)$ допускает *касательные миноранты* $\varphi(x,y,\theta)$ в точках $y\in X$, т.е. почти для всех θ выполнено

1) $f(x,\theta)\ge\varphi(x,y,\theta)$ для всех $x\in X$, $y\in X$;
2) $f(y,\theta)=\varphi(y,y,\theta)$ для всех $y\in X$;
3) функция $\varphi(x,y,\theta)$ непрерывна по (x,y) почти для всех θ, $y\in X$;
4) $\varphi(x,y,\theta)$ - измерима по θ для любых $x,y\in X$;
5) $|\varphi(x,y,\theta)|\le M(\theta)$ для всех $x,y\in X$ с интегрируемой функцией $M(\theta)$.

Тогда функции $\varphi(x,y)=E\varphi(x,y,\theta)$ являются касательными минорантами для функции математического ожидания $F(x)=Ef(x,\theta)$.

Доказательство. Условия (i), (ii) определения 1 следуют из 1), 2). Условие (iii) следует из 3), 4) по теореме Лебега о мажорируемой сходимости и из леммы 1.

Примечание. Лемма 2 дает способ построения (стохастических) касательных минорант для функций типа математического ожидания. Касательные миноранты функции вероятности $P(x)=P\{f(x,\theta)\ge 0\}$ строятся аналогично, а именно, в качестве касательной в точке y миноранты $P(x)$ можно взять функцию $\phi(x,y)=P\{\varphi(x,y,\theta)\ge 0\}$, где $\varphi(x,y,\theta)$ - касательная в точке y миноранта функции $f(x,\theta)$.

Если функции $f(x,\theta)$ Липшицевы (Гельдеровы) с интегрируемой по θ константой Липшица $L(\theta)$ и показателем α, то в качестве касательной в точке y миноранты для $f(x,\theta)$ можно взять функцию

$$\varphi(x,y,\theta) = f(y,\theta) - L(\theta)\|x-y\|^{\alpha}.$$

Для гладких по x функций $f(x,\theta)$ с липшицевым градиентом (с константой $L(\theta)$) в качестве стохастических касательных минорант можно использовать касательные к графику $f(\cdot,\theta)$ в точках y параболоиды:

$$\varphi(x,y,\theta) = f(y,\theta) + \frac{1}{2L(\theta)}\|\nabla f(y,\theta)\|^2 - \frac{L(\theta)}{2}\left\|x - y - \frac{1}{L(\theta)}\nabla f(y,\theta)\right\|^2.$$

О некоторых модификациях метода Пиявского. Ограничимся случаем задачи без общих ограничений:

$$F(x) \to \min\nolimits_{x\in X}. \qquad (2)$$

Предположим, что функция $F(x)$ допускает касательные в точке $y \in X$ миноранты $\varphi(x,y)$.

Алгоритм 1. Точки $y^0,\ldots,y^{k_0} \in X$ произвольны. Пусть уже построены точки $y^0,\ldots,y^k \in X$. Точку y^{k+1}, $k \ge k_0$, найдем как решение следующей специальной многоэкстремальной задачи:

$$\varphi_k(x) := \max\{\varphi_{k-1}(x), \varphi(x,y^k) - t_k(F(y^k) - \varphi_{k-1}(y^k)\} \to \min\nolimits_{x\in X}, \qquad (3)$$

где $0 \le t_k \le \overline{t} < 1$. Задача (3) является задачей параметрического программирования с параметром $t = t_k$. При $t_k \equiv 0$ получаем стандартный метод Пиявского. Доказательство сходимости алгоритма в основном следует [5].

Лемма 3 [5]. Пусть функции $\phi_k(y)$ определены в некоторой окрестности $V(y')$ точки y', непрерывны в точке y' равностепенно по k и существует конечный предел $\lim_{k\to\infty}\phi_k(x) = \phi(x)$ для $x \in V(y')$. Тогда для любой последовательности $\{y^k\}$, сходящейся к точке y', имеет место

$$\lim\nolimits_{k\to\infty}\phi_k(y^k) = \lim\nolimits_{k\to\infty}\phi(y^k) = \phi(y').$$

Лемма 4. Имеет место $\lim_{k\to\infty}\left[F(y^k) - \varphi_{k-1}(y^k)\right] = 0$.

Доказательство. Предположим противное. Тогда существует последовательность $\left\{y^{k_m}\right\}$ такая, что $\lim_{k\to\infty} y^{k_m} = y$ и

$$\lim\nolimits_{m\to\infty}\left[F(y^{k_m}) - \varphi_{k-1}(y^{k_m})\right] = \varepsilon > 0.$$

Последовательности $\{\varphi_k(x)\}$, $x \in X$, монотонно возрастают и ограничены сверху функцией $F(x)$, поэтому существует предел $\varphi(x) = \lim_{k\to\infty}\varphi_k(x)$. Функции $\varphi(x,y^k)$, а следом и $\varphi_k(x)$ равностепенно непрерывны в каждой точке, поэтому согласно лемме 3 $\lim_{m\to\infty}\varphi_{k_m}(y^{k_m}) = \varphi(y)$, $\lim_{m\to\infty}\varphi(y^{k_m}) = \varphi(y)$. Тогда

$$\lim_{m\to\infty}\left[F(y^{k_m})-\varphi_{k-1}(y^{k_m})\right]=F(y)-\varphi(y)=\varepsilon>0. \qquad (4)$$

Но с другой стороны

$$\varphi_k(y^{k_m})=\max\{\varphi_{k_m-1}(y^{k_m}),\varphi(y^{k_m},y^{k_m})-t_{k_m}(F(y^{k_m})-\varphi_{k_m-1}(y^{k_m})\}=$$
$$=\max\{\varphi_{k_m-1}(y^{k_m}),f(y^{k_m})-t_{k_m}(F(y^{k_m})-\varphi_{k_m-1}(y^{k_m})\}=$$
$$=\max\{\varphi_{k_m-1}(y^{k_m}),(1-t_{k_m})F(y^{k_m})+t_{k_m}\varphi_{k_m-1}(y^{k_m})\}=$$
$$=F(y^{k_m})-t_{k_m}(F(y^{k_m})-\varphi_{k_m-1}(y^{k_m}),$$

откуда

$$t_{k_m}(F(y^{k_m})-\varphi_{k_m-1}(y^{k_m}))=F(y^{k_m})-\varphi_{k_m}(y^{k_m}) \qquad (5)$$

В силу леммы 3 $\lim_{m\to\infty}\varphi_{k_m}(y^{k_m})=\varphi(y)$, $\lim_{m\to\infty}\varphi_{k_m-1}(y^{k_m})=\varphi(y)$, а в силу непрерывности $F(x)$ имеет место $\lim_{m\to\infty}F(y^{k_m})=F(y)$. Переходя в (5) к пределу по $m\to\infty$, получаем

$$(F(y)-\varphi(y))\overline{\lim}_{m\to\infty}t_{k_m}\geq F(y)-\varphi(y).$$

В силу (4) $F(y)-\varphi(y)\neq 0$, поэтому $\overline{\lim}_{m\to\infty}t_{k_m}\geq 1$, что противоречит условию $t_k\leq\overline{t}<1$. Лемма доказана.

Теорема 1. Все предельные точки последовательности $\left\{y^k\right\}$ являются точками глобального минимума задачи (2).

Доказательство аналогично доказательству теоремы 3 из [5].

Алгоритм 2. Точки $y^0,\dots,y^{k_0}\in X$ произвольны. Пусть уже построены точки $y^0,\dots,y^k\in X$. Обозначим $\phi^k(x)=\max_{0\leq i\leq k}\varphi(x,y^i)$. Точку $y^{k+1},k\geq k_0$, найдем как решение следующей параметрической специальной многоэкстремальной задачи:

$$\varphi_k(x):=(1-t_k)\varphi_{k-1}(x)+t_k\phi^k(x)\to\min_{x\in X}, \qquad (6)$$

где $0<\underline{t}\leq t_k\leq 1$. Задача (2) является задачей параметрического программирования с параметром $t=t_k$. При $t_k\equiv 1$ получаем стандартный метод Пиявского.

Доказательство сходимости алгоритма 2 аналогично доказательству сходимости алгоритма 1.

Метод стохастических минорант (стохастический аналог метода Пиявского). Аппроксимируем задачу (1) эмпирическими средними:

$$\min_{x\in X}[F_N(x)=\frac{1}{N}\sum_{k=1}^N f(x,\theta^k)], \qquad (7)$$

где θ^k – независимые наблюдения случайного параметра θ. Если функции $F_N(x)$ равномерно сходятся к $F(x) = Ef(x,\theta)$, то вместо исходной задачи (1) можно решать приближенную задачу (7).

Примечание. Если $|f(x,\theta)| \le M(\theta)$ для любого $x \in X$ с интегрируемой функцией $M(\theta)$, то семейство функций $\{f(x,\cdot), x \in X\}$ равномерно интегрируемо и с вероятностью 1 $F_N(x)$ равномерно на X сходятся к функции $F(x) = Ef(x,\theta)$ [7].

Функции $\varphi_N(x,y) = (1/N)\sum_{k=1}^{N} \varphi(x,y,\theta^k)$, очевидно, являются касательными минорантами для $F_N(x)$. Тогда для решения приближенной детерминированной задачи (7) применим метод Пиявского, например, алгоритмы 1, 2.

Численные эксперименты. Для экспериментов использовались функции:

1) $F_1(x) = 10 + x^2 - 10\cos(2\pi x)$, $x \in [-5;5]$, с оценками констант Липшица функции $l_1 = 72$ и ее градиента $L_1 = 397$, и с глобальным минимумом в точке $x^* = 0$ и минимальным значением $f(x^*) = 0$; **2)** $F_2(x) = 418.9829 - x\sin(\sqrt{x})$, $x \in [5;500]$; $l_2 = 12$; $L_2 = 0.41$; $x^* = 420.9687$, $f(x^*) = 0$.

В таблицах 1, 2 приведено число итераций метода Пиявского с использованием касательных конусов и параболоидов (парабол) для глобальной минимизации функций F_1, F_2. Вариант метода с использованием парабол демонстрирует гораздо большую эффективность.

ТАБЛИЦА 1. Число итераций для заданной точности минимизации F_1

Точность	10^{-4}	10^{-3}	10^{-2}	10^{-1}
Параболоиды	47	43	39	37
Конусы	1390	451	198	92

ТАБЛИЦА 2. Число итераций для заданной точности минимизации F_2

Точность	10^{-4}	10^{-3}	10^{-2}	10^{-1}
Параболоиды	29	24	23	19
Конусы	9264	3202	822	290

В таблице 3 продемонстрирована работа модифицированного алгоритма 1 с использованием касательных параболоидов на функциях F_1, F_2 соответственно, $t_k \equiv 0.3$. Наблюдается интересный эффект отбрасывания (полного мажорирования) части минорант (полное число минорант равно числу итераций).

ТАБЛИЦА 3. Результаты минимизации F_1 модифицированным алгоритмом 1

Номер итерации	Количество минорант	Точность	Номер итерации	Количество минорант	Точность
10	8	543,44	70	38	28,14
20	12	122,01	80	43	7,49
30	19	59,53	90	48	0,26
40	23	23,34	100	52	0,0008
50	27	13,26	105	56	0,00006
60	33	37,51			

Рассмотрим пример задачи (стохастической) глобальной минимизации функции $F(x) = Ef(x,\theta)$, где $f(x,\theta) = x^4 - 13x^2 - 6x - 14x^2\theta - 15x\theta^2$, θ – равномерно распределенная на отрезке [0;1] случайная величина, $x \in [-4;4]$. Нетрудно вычислить, что $F(x) = Ef(x,\theta) = x^4 - 20x^2 - 11x$, $x^* = 3.29$; $F(x^*) = -135.51$. Возьмем детерминированные оценки констант Липшица $l = 107$ и $L = 152$ функции $f(x,\theta)$ и ее градиента соответственно. Аппроксимируем $F(x)$ эмпирической функцией $F_N(x) = (1/N)\sum_{i=1}^{N} f(x,\theta_i)$, где θ_i – независимые равномерно распределенные на отрезке [0;1] случайные величины. Результаты минимизации $F_N(x)$ методом Пиявского с использованием параболических и конусных минорант для различных N представлены в табл. 4.

ТАБЛИЦА 4. Результаты минимизации функций F_N для различных N

N	Параболы			Конусы		
	Количество итераций	x^*	$F_N(x^*)$	Количество итераций	x^*	$F_N(x^*)$
10	19	3,26	-131,46	472	3,20	122,55
20	16	3,36	-144,07	369	3,46	-167,13
50	18	3,26	-132,26	358	3,23	-125,23
100	17	3,26	-130,13	386	3,32	-141,71
200	18	3,28	-130,68	366	3,29	-135,89
400	18	3,29	-136,14	347	3,28	-134,41

Заключение. В работе обоснован метод Пиявского для решения задачи стохастической глобальной оптимизации с целевой функцией типа математического ожидания. В частности, указан способ вычисления касательных минорант для таких функций, а именно, в качестве минорант функции математического ожидания можно брать математическое ожидание стохастических касательных минорант подынтегральной функции. Здесь ситуация подобна той, что возникает при вычислении градиентов интегральных функционалов. Вычислить градиент или миноранту интегрального функционала может быть весьма трудно, а вычислить (стохастический) градиент или (стохастическую) миноранту подынтегральной функции может быть сравнительно легко. Предлагается аппроксимировать исходную целевую функцию ее эмпирической оценкой. Знание стохастических касательных минорант позволяет легко построить касательные миноранты для оценок и, таким образом применять метод Пиявского для глобальной минимизации аппроксимаций. В работе также обоснованы модификации исходного метода Пиявского, связанные с использованием не касающихся минорант. На численных примерах показано, что использование касательных параболоидов (в одномерном случае парабол) значительно повышает эффективность метода по сравнению с классическим вариантом, использующим касательные конусы.

В.І. Норкін, Б.О. Онищенко

ПРО СТОХАСТИЧНИЙ АНАЛОГ МЕТОДУ ГЛОБАЛЬНОЇ ОПТИМІЗАЦІЇ ПІЯВСЬКОГО

У роботі обґрунтований метод Піявського для розв'язування задачі стохастичної глобальної оптимізації з цільовою функцією типу математичного очікування. Зокрема, показано, що в якості дотичних мінорант для таких функцій можна брати математичне очікування стохастичних дотичних мінорант підінтегральної функції. Обґрунтовані модифікації методу, пов'язані з використанням не дотичних мінорант.

V.I. Norkin, B.O. Onischenko

ON A STOCHASTIC ANALOGUE OF PIYAVSKII's GLOBAL OPTIMIZATION METHOD

In the paper Piyavskii's global optimization method is validated for stochastic global optimization problem with objective function in the form of mathematical expectation. In particular, it is shown that as a tangent minorant of such function one can take a mathematical expectation of the stochastic tangent minorant of the underintegral function. Modifications of the method, using nontangent minorants, are considered.

1. *Пиявский С.А.* Алгоритм отыскания абосолютного минимума функций // Теория оптимальных решений. – Киев: Ин-т кибернетики АН УССР, 1967. – Вып. 2. – С. 13–24.
2. *Данилин Ю.М., Пиявский А.С.* Об одном алгоритме отыскания абсолютного минимума // Теория оптимальных решений. – Киев: Ин-т кибернетики АН УССР, 1967. – Вып. 2. – С. 25–37.
3. *Пиявский С.А.* Один алгоритм отыскания абосолютного минимума функций // Журнал вычислительной математики и математической физики – 1972. – 12, № 4. – С. 888–896.
4. *Horst R., Tuy H.* Global Optimization (Deterministic Approaches). 3rd, revised and enlarged edition. – Berlin: Springer Verlag, 1996. – 600 p.
5. *Норкин В.И.* О методе Пиявского для решения общей задачи глобальной оптимизации // Журнал вычислительной математики и математической физики – 1992. – 32, № 7. – С. 992–1007.
6. *Khamisov O.* On Optimization Properties of Functions, with a Concave Minorant // J. Of Global Optimization. – 1999. – 14, № 1. – P. 79–101.
7. *Le Cam L.* On some asymptotic properties of maximum likelihood estimates and related Bayes' estimated // Univ. California Publ. Statist. – 1953. – 1. – P. 227–330.

Получено 03.07.2003