О МИНОРАНТНЫХ МЕТОДАХ СТОХАСТИЧЕСКОЙ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

Онищенко Б.О.
Институт кибернетики НАНУ
boris@cdu.edu.ua

Содержание

- **Задача стохастической глобальной оптимизации.**
- **Стохастические касательные миноранты.**
- **Способы построения стохастических касательных минорант.**
- **Метод последовательных эмпирических аппроксимаций.**
- **Модификации метода Пиявского.**
- **Стохастический метод ветвей и границ.**
- **Численные эксперименты.**

Литература:

– *Норкин В.И., Онищенко Б.О.* О стохастическом аналоге метода глобальной оптимизации Пиявского // Теория оптимальных решений. – Киев: Ин-т кибернетики АН Украины, 2003.
– *Норкин В.И., Онищенко Б.О.* Метод ветвей и границ с минорантными оценками для решения задач стохастической глобальной оптимизации // Компьютерная математика. – Киев: Ин-т кибернетики АН Украины, 2003.
– *Пиявский* (1967); *Даналин, Пиявский* (1967); *Норкин* (1992); *Horst, Tuy* (1996).

Задача стохастической глобальной оптимизации.

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, $\theta \in \Theta$, (Θ, Σ, P) - вероятностное пространство задачи, X - непрерывное или дискретное множество.

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

i) $F(x) \geq \varphi(x,y)$ 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 .

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

Стохастические касательные 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 минорантами для $F(x)$.

Лемма 1. 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) $f(x, \theta) \geq \varphi(x, y, \theta)$ 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;

4) $\varphi(x, y, \theta)$ - измерима по θ для любых $x, y \in X$;

5) $|\varphi(x, y, \theta)| \leq M(\theta)$ для всех $x, y \in X$ с некоторой интегрируемой функцией $M(\theta)$.

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

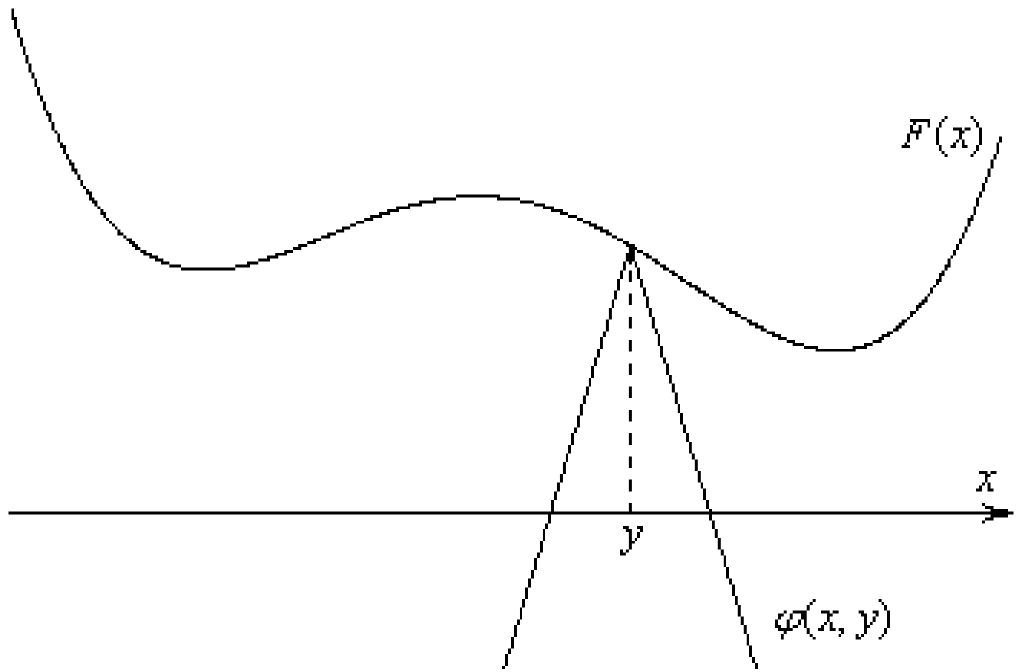
Способы построения стохастических касательных минорант (касательные конусы).

1. Если функции $f(x,\theta)$ липшицевы (гельдеровы) с интегрируемой по θ константой Липшица $L(\theta)$ и показателем α:

$$\left| f(x^1,\theta) - f(x^2,\theta) \right| \le L(\theta) \left\| x^1 - x^2 \right\|^{\alpha},\ \forall\, x^1, x^2 \in X,\ 0 < \alpha \le 1,$$

то в качестве касательной в точке y миноранти для $f(x,\theta)$ можно взять функцию

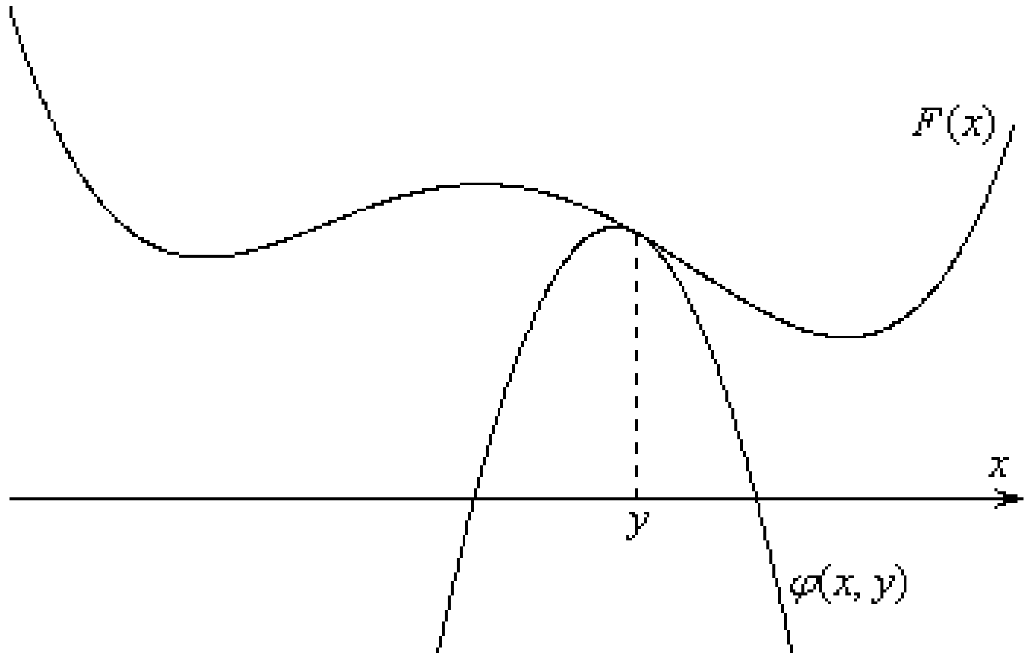
$$\varphi(x, y, \theta) = f(y,\theta) - L(\theta) \left\| x - y \right\|^{\alpha}.$$



Способы построения стохастических касательных минорант (касательные параболоиды).

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

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



В этих примерах возможны различные нормы, например, для n-мерного вектора z можно использовать норму $\|z\| = \left(\sum_{i=1}^{n} z_i^{\,2}\right)^{1/2}$ или $\|z\| = \max\limits_{1\le i\le n}|z_i|$.

Метод последовательных эмпирических аппроксимаций.

Аппроксимируем задачу (1) эмпирическими средними:

$$\min_{x\in X}[F_N(x) = (1/N)\sum_{k=1}^{N} f(x,\theta^k)], \tag{2}$$

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 (2). Следующее утверждение обосновывают такой подход.

Теорема 1 (LeCam (1953)). Пусть функция $f(x,\theta)$ непрерывна по x для почти всех $\theta \in \Theta$ и измерима по θ для всех $x \in X$, X - компакт в R^n. Предположим, что семейство $\{f(x,\cdot), x \in X\}$ равномерно интегрируемо. Тогда с вероятностью 1 функции $F_N(x)$ равномерно на X сходятся к $F(x) = Ef(x,\theta)$.

Функции $\varphi_N(x,y) = (1/N)\sum_{k=1}^{N}\varphi(x,y,\theta^k)$, очевидно, являются ***касательными минорантами*** для $F_N(x)$.

Важно заметить, что в данном методе можно использовать как ***фиксированные*** значения N, так и постепенно ***наращивать*** его значение (т.е. $N \to \infty$).

Модификации метода Пиявского.

Пусть имеем задачу без общих ограничений:

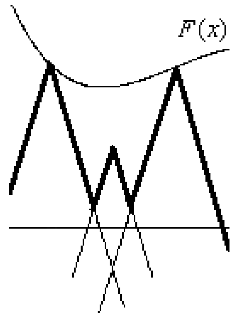
$$F(x) \to \min_{x \in X}. \qquad (3)$$

Точка $y^0 \in X$ произвольна; $\varphi_0(x,t) = \varphi(x, y^0)$. Пусть уже построены точки $y^1, ..., y^k \in X$.

1) Классический метод Пиявского:

$$\varphi_k(x) = \max\{\varphi_{k-1}(x), \varphi(x, y^k)\}; \qquad (4)$$

2) Модификация 1:



$$\varphi_k(x,t) = \max\{\varphi_{k-1}(x), \varphi(x, y^k) - t(F(y^k) - \varphi_{k-1}(y^k))\}, \qquad (5)$$

где $0 \le t \le \overline{t} < 1$. Здесь не меняется форма составляющих минорант;

3) Модификация 2:

$$\phi^k(x) = \max_{0 \le i \le k} \varphi(x, y^i),$$

$$\varphi_k(x,t) = (1-t)\varphi_{k-1}(x) + t\phi^k(x), \qquad (6)$$

где $0 < \underline{t} \le t \le 1$. Здесь используемые миноранты деформируются.

Точку $y^{k+1}, k \ge 0$, найдем как решение специальной многоэкстремальной задачи:

$$\varphi_k(x) := \varphi_k(x,t) \to \min_{x \in X}. \qquad (7)$$

Очевидно, что обе модификации метода в общем случае используют **не касающиеся миноранты**.

Все модификации, как и классический вариант метода, можно использовать для решения **стохастических задач глобальной оптимизации**, используя при этом **метод последовательных эмпирических аппроксимаций**.

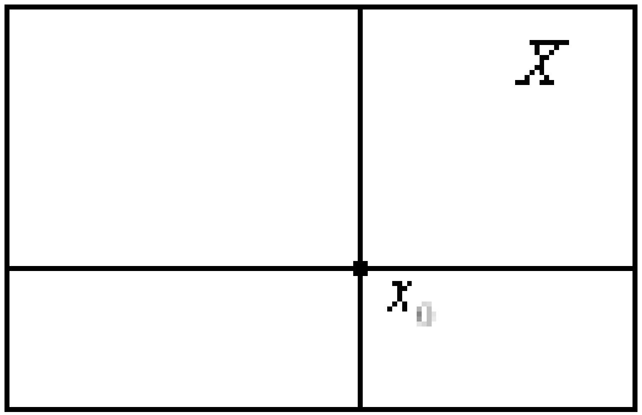
<u>Стохастический метод ветвей и границ.</u>

Пусть имеем задачу без общих ограничений:

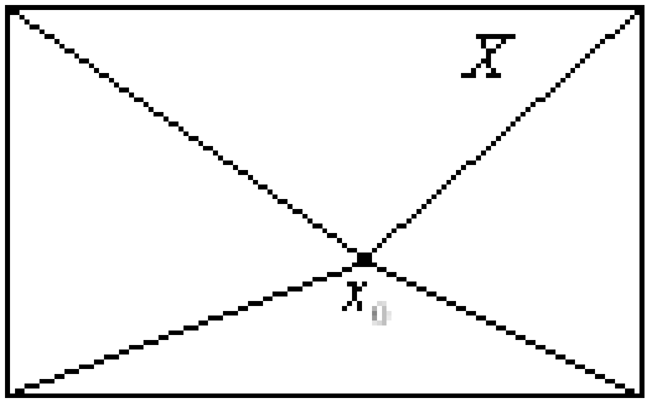
$$F(x) \to \min_{x \in X},\ X = [a,b]^N. \qquad (8)$$

Способы начального разбиения множества X:

1) <u>Параллелепипедами;</u>



2) <u>Симплексами.</u>

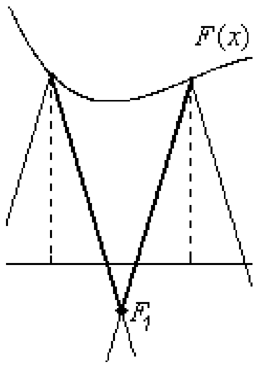


Пусть $\{\varphi(x,y)\}_{y \in X}$ - семейство касательных в точках $y \in X$ минорант для $F(x)$, $\{y \in Z \subseteq X\}$ – конечное или бесконечное множество точек из X.

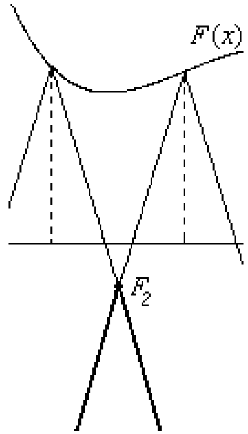
<u>Стохастический метод ветвей и границ.</u>

Минорантные границы оптимальных значений целевой функции на фрагменте разбиения.

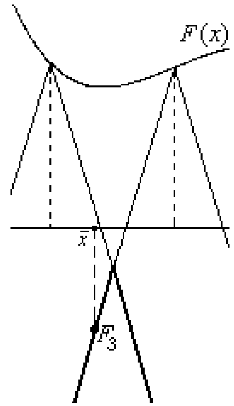
<u>Рассмотрим некоторые оценки снизу для функции</u> $F(x)$:



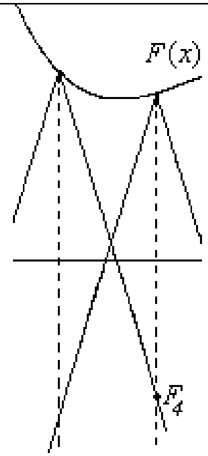
$F_1 = \min_{x\in X} \max_{y\in Z} \varphi(x, y)$



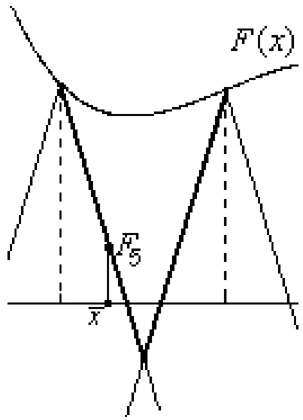
$F_2 = \max_{x\in X} \min_{y\in Z} \varphi(x, y)$



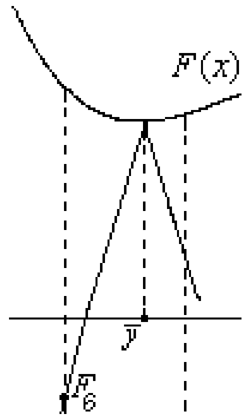
$F_3 = \min_{y\in Z} \varphi(\overline{x}, y)$



$F_4 = \max_{y\in Z} \min_{x\in X} \varphi(x, y)$



$F_5 = \max_{y\in Z} \varphi(\overline{x}, y)$



$F_6 = \min_{x\in Z}\{\varphi(x, \overline{y})\}$

<u>Стохастический метод ветвей и границ.</u>

Специфические минорантные границы оптимальных значений для функции математического ожидания.

Если учесть, что $F(x) = Ef(x,\theta)$ и $\varphi(x,y) = E\varphi(x,y,\theta)$, то можно построить и специфические оценки снизу для функции $F(x)$:

- $\Phi_1 = \min_{x\in X} E\max_{y\in X} \varphi(x,y,\theta)$;
- $\Phi_2 = E\min_{x\in X} \max_{y\in X} \varphi(x,y,\theta)$;
- $\Phi_3 = E\max_{y\in X} \min_{x\in X} \varphi(x,y,\theta)$;
- $\Phi_4 = E\max_{x\in X} \min_{y\in X} \varphi(x,y,\theta)$;
- $\Phi_5 = E\min_{y\in X} \varphi(\overline{x},y,\theta)$, $\overline{x} \in X$.

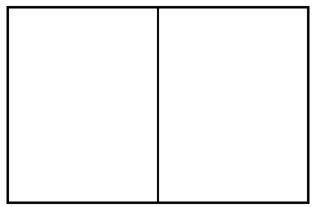
Как видно из представленных оценок, если целевая функция является функцией математического ожидания некоторой функции, то возможно внесение операций взятия максимума или минимума под знак математического ожидания.

Стохастический метод ветвей и границ.

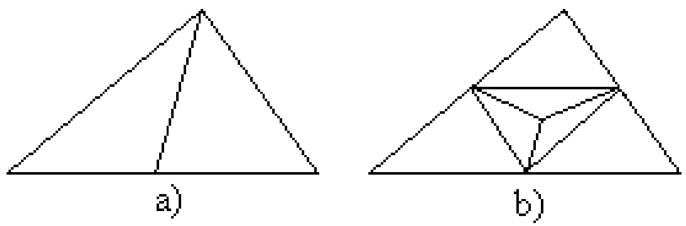
Способы размельчения фрагмента разбиения, имеющего наименьшую оценку снизу.

При размельчении «рекордного» фрагмента разбиения необходимо, чтобы его диаметр стремился к нулю. Для выполнения данного условия можно использовать следующие способы размельчения:

1) начальное разбиение параллелепипедами – ребра наибольшей длины параллелепипеда делятся пополам, и он разбивается на два новых параллелепипеда;



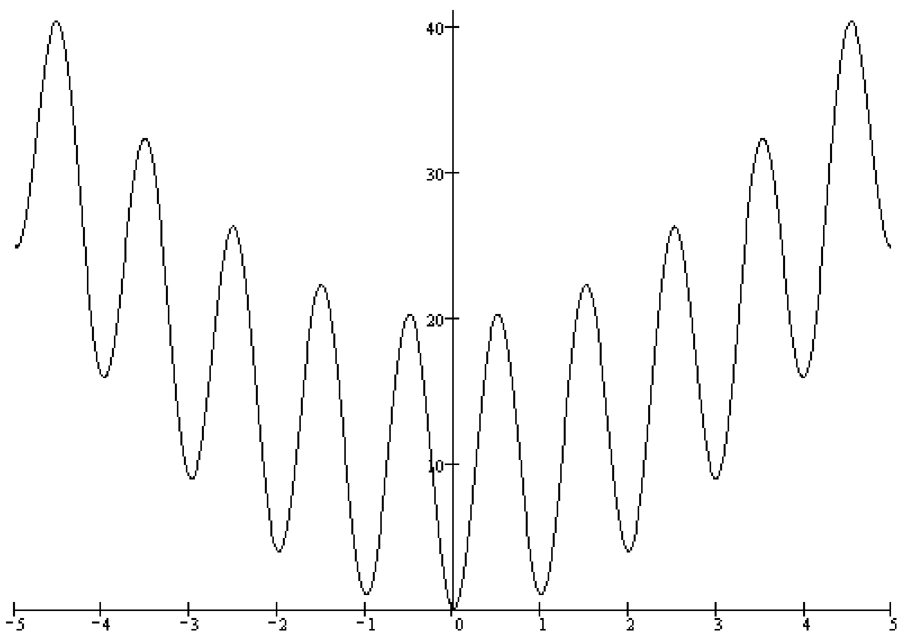
2) начальное разбиение симплексами:
   a) ребро наибольшей длины симплекса делятся пополам, и он разбивается на два новых симплекса;
   b) все ребра симплекса делятся пополам, также находится центр тяжести средин ребер, и он разбивается на новые симплексы с участием всех полученных точек.



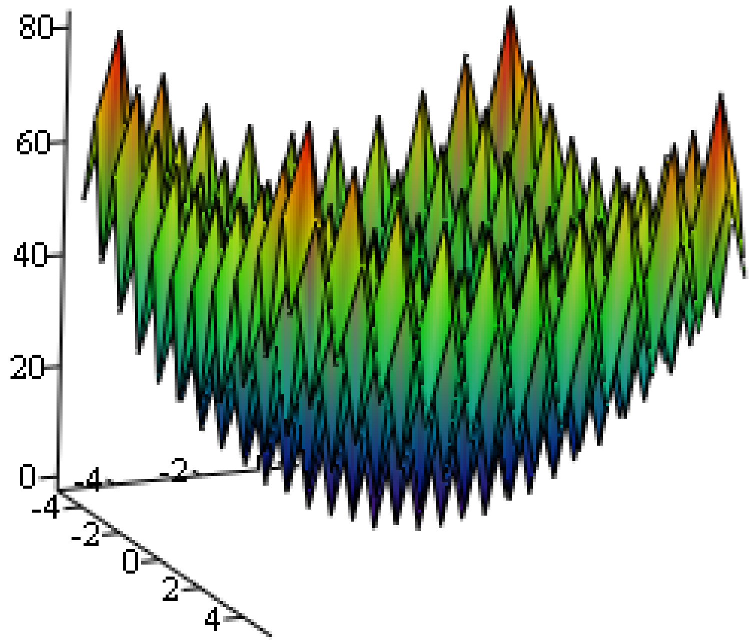
Численные эксперименты.

Тестовая функция

Для экспериментов использовалась следующая тестовая функция: $f(x) = 10n + \sum_{i=1}^{n} (x_i^2 - 10\cos(2\pi x_i))$, где n – размерность пространства $x \in [-5;5]$, с оценками констант Липшица функции $l = 72$ и ее градиента $L = 397$; $x^* = 0$, $f(x^*) = 0$;



$n = 1$



$n = 2$

<u>Численные эксперименты.</u>

Модификации метода Пиявского

ТАБЛИЦА 1. Результаты работы классического метода Пиявского (4).

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

Таблица 1 показывает, что использование касательных параболоидов ведет к существенному сокращению количества итераций, необходимых для достижения заданной точности.

ТАБЛИЦА 2. Результаты работы модифицированного алгоритма Пиявского (5).

Номер итерации	К-во минорант	Точность	Номер итерации	К-во 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			

Таблица 2 показывает, что использование не касающихся минорант приводит к отбрасыванию (полному мажорированию) некоторых составляющих минорант, что в свою очередь упрощает поиск минимума результирующей миноранты.

Численные эксперименты.

Метод ветвей и границ

ТАБЛИЦА 3. Результаты работы алгоритма метода ветвей и границ.

Размерность пространства	Способ разбиения	Вид минорант	Характеристики алгоритма	Вид оценки снизу			
				F_3	F_4	F_5	F_6
1	1, 2a, 2b	Параболоид	Итерации	47	67	25	59
			Оставленные симплексы	3	2	3	3
		Конус	Итерации	197	293	93	163
			Оставленные симплексы	71	141	58	75
2	1	Параболоид	Итерации	1245	3184	618	1612
			Оставленные симплексы	18	41	7	21
	2a		Итерации	2555	4531	517	2847
			Оставленные симплексы	27	55	8	49
	2b		Итерации	2271	3707	269	2661
			Оставленные симплексы	22	67	17	69

Таблица 3 показывает, что наиболее эффективно метод работает при использовании эвристической оценки снизу $F_5 = \max_{y \in Z} \varphi(\overline{x}, y)$.

Численные эксперименты.

Стохастическая глобальная оптимизация

Для тестирования стохастических алгоритмов использовалась функция:

$$f(x,\theta)=(x-1-\theta)(x-3-\theta)(x-7-\theta)(x-11-\theta),$$

где θ – равномерно распределенная на отрезке $[0;1]$ случайная величина, $x\in[1;11]$. Нетрудно вычислить, что

$$F(x)=Ef(x,\theta)=x^4-24x^3+187x^2-537x+\frac{14051}{30}.$$

Детерминированные оценки констант Липшица $l=537$, $L=374$ функции $f(x,\theta)$ и ее градиента. $x^*=9.977$; $F(x^*)=-201.667$.

Аппроксимируем $F(x)$ **эмпирической функцией** $F_N(x)=\left(1/N\right)\sum_{i=1}^{N} f(x,\theta_i)$, где θ_i – независимые равномерно распределенные на отрезке $[0;1]$ случайные величины.

ТАБЛИЦА 4. Результаты работы алгоритма метода ветвей и границ при $N=300$.

Вид минорант	Х-ки	Вид оценки снизу					
		F_3	F_4	Φ_3	Φ_5	F_5	F_6
Параболоид	Итерации	20	32	31	20	19	26
Конус	Итерации	1991	3935	3886	2041	1961	1973

Таблица 4 показывает, что оценки использующие внесение операции взятия максимума или минимума под знак математического ожидания работают не менее эффективно, чем оценки самой функции математического ожидания.

<u>Численные эксперименты.</u>

Стохастическая глобальная оптимизация

ТАБЛИЦА 5. Оценка работы алгоритма, при N пропорциональном числу итераций.

Итерации	F_3		F_4		Φ_3	
	Алгоритмическая точность	Точность	Алгоритмическая точность	Точность	Алгоритмическая точность	Точность
10	22.645	11.656	321.518	186.017	321.518	186.017
20	0.095	2.091	4.221	0.842	3.423	0.842
50	0.049	0.715	0.056	0.709	0.018	0.678
100	0.001	0.452	0.000	0.452	0.006	0.445
200	0.005	0.412	0.069	0.359	0.055	0.343
300	0.009	0.335	0.009	0.335	0.011	0.300
	Φ_5		F_5		F_6	
10	22.645	11.656	6.426	1.920	35.065	11.656
20	0.095	2.091	0.117	2.093	0.057	2.007
50	0.049	0.715	0.049	0.715	0.054	0.709
100	0.001	0.452	0.001	0.452	0.001	0.452
200	0.073	0.356	0.073	0.356	0.068	0.359
300	0.009	0.335	0.009	0.335	0.009	0.335

Таблица 5 показывает, что в стохастическом методе ветвей и границ при использовании последовательных эмпирических аппроксимаций возможно использование не фиксированного числа наблюдений случайной величины, а постепенно увеличивать его с ростом числа итераций.