ОБ ОБОБЩЕНИИ МИНОРАНТНЫХ МЕТОДОВ СТОХАСТИЧЕСКОЙ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

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

Содержание

- **Постанова задачи.**
- **Стохастические касательные миноранты.**
- **Метод последовательных эмпирических аппроксимаций.**
- **Модификации метода Пиявского.**
- **Стохастический метод ветвей и границ.**
- **Метод ветвей и приближенных границ.**
- **Численные эксперименты.**

Литература:

– *Норкин В.И., Онищенко Б.О.* О стохастическом аналоге метода глобальной оптимизации Пиявского // Теория оптимальных решений. – К.: Ин. киб. АНУ, 2003.
– *Норкин В.И., Онищенко Б.О.* Метод ветвей и границ с минорантными оценками для решения задач стохастической глобальной оптимизации // Компьютерная математика. – К.: Ин. киб. АНУ, 2003.
– *Норкин В.И., Онищенко Б.О.* О глобальной минимизации функций минимума методом минорант // Теория оптимальных решений. – К.: Ин. киб. АНУ, 2004
– *Норкин В.И., Онищенко Б.О.* О методе ветвей и приближенных границ // Компьютерная математика. – Киев: Ин. киб. АНУ, 2004.
– *Пиявский* (1967); *Даналин, Пиявский* (1967); *Норкин* (1992); *Horst, Tuy* (1996).

Постановка задачи.

Рассмотрим следующие задачи **стохастической глобальной оптимизации**:

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

$$\min_{x\in X}[F(x) = E_x f(x,\theta_x)], \qquad (2)$$

где θ - случайный параметр, E – символ математического ожидания по θ, E_x– символ математического ожидания по θ_x, значения которого зависят от x, $f(x,\theta)$ – 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 $F(x) = Ef(x, \theta)$.

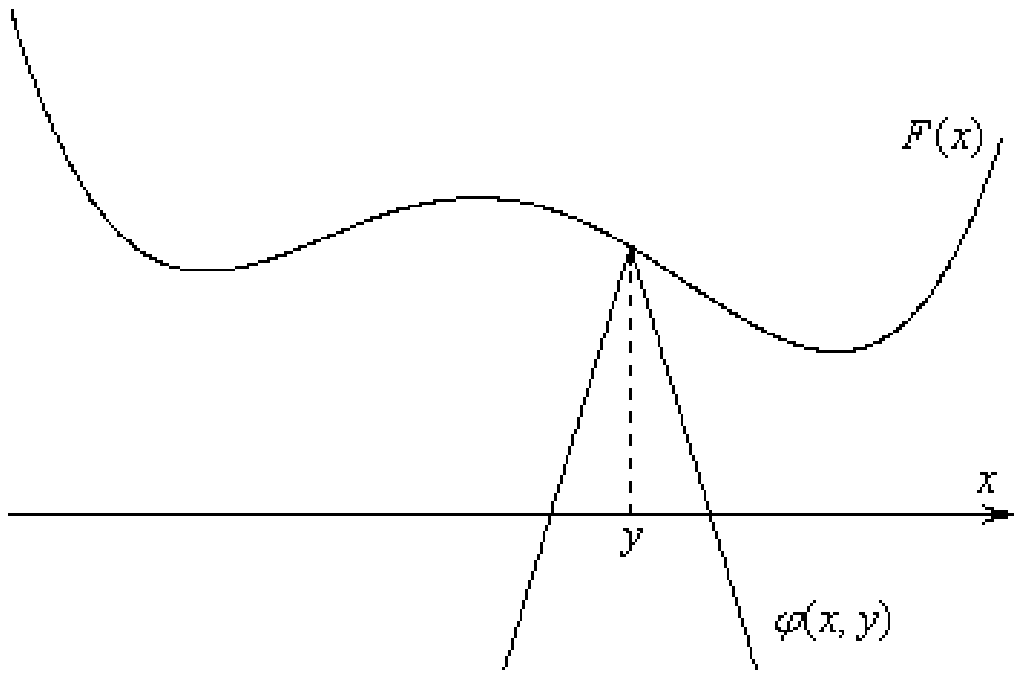
Способы построения стохастических касательных минорант (касательные конусы).

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

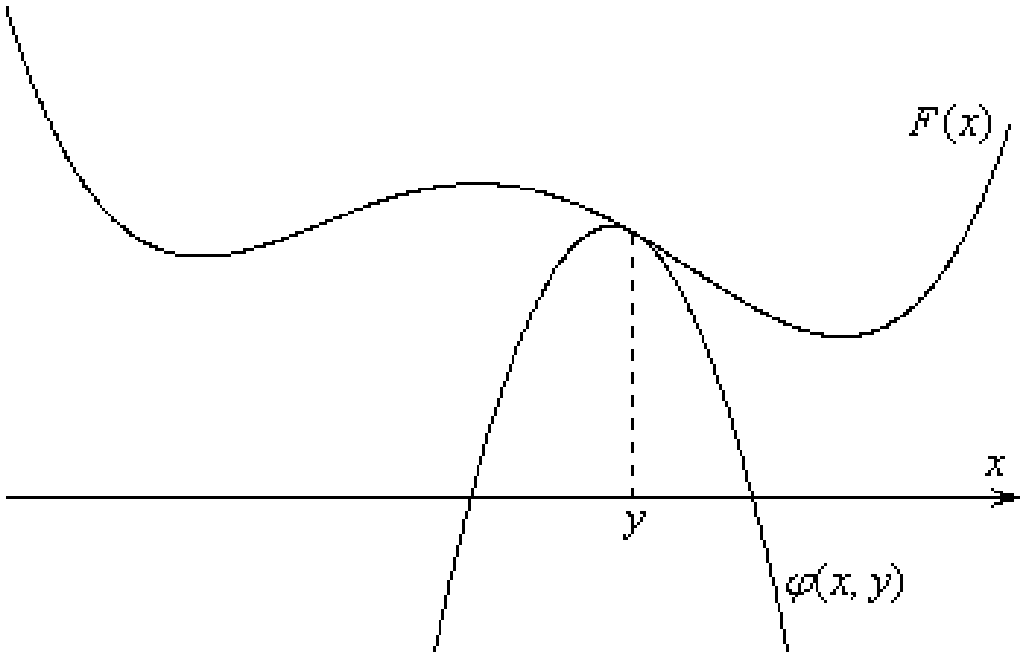
$$\varphi(x,y,\theta) = f(y,\theta) - L(\theta) \left\| x - y \right\|^{\alpha}.$$



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 $\|z\| = \max\limits_{1\le i\le n}|z_i|$.

Способы построения стохастических касательных минорант (разность выпуклых функций).

3. Если функция $f(x,\theta)$ представима в виде разности выпуклых функций $f_1(x,\theta)$ и $f_2(x,\theta)$

$$f(x,\theta) = f_1(x,\theta) - f_2(x,\theta),$$

то касательная в точке y минората к $f(x,\theta)$ может быть построена в виде:

$$\varphi(x,y,\theta) = l_1(x,y,\theta) - f_2(x,y,\theta),$$

где $l_1(x,y,\theta)$ – касательная в точке y гиперплоскость к графику функции $f_1(x,\theta)$.

<u>Метод последовательных эмпирических аппроксимаций.</u>

Аппроксимируем задачу (1) эмпирическими средними:

$$\min_{x\in X}[F_N(x)=\left(1/N\right)\sum_{k=1}^{N}f(x,\theta^k)], \qquad (3)$$

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 (3). 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 $\left\{f(x,\cdot), x\in X\right\}$ равномерно интегрируемо. Тогда с вероятностью 1 функции $F_N(x)$ равномерно на X сходятся к $F(x)=Ef(x,\theta)$.

Функции $\varphi_N(x,y)=\left(1/N\right)\sum_{k=1}^{N}\varphi(x,y,\theta^k)$, 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:

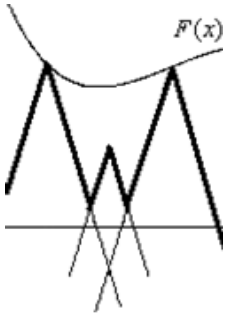
$$F(x) \to \min_{x \in X}. \tag{4}$$

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:

$$\varphi_k(x) = \max\{\varphi_{k-1}(x), \varphi(x, y^k)\}; \tag{5}$$

2) Модификация 1:



$$\varphi_k(x,t) = \max\{\varphi_{k-1}(x), \varphi(x, y^k) - t(F(y^k) - \varphi_{k-1}(y^k)\}, \tag{6}$$

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

3) Модификация 2:

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

$$\varphi_k(x,t) = (1-t)\varphi_{k-1}(x) + t\phi^k(x), \tag{7}$$

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:

$$\varphi_k(x) := \varphi_k(x,t) \to \min_{x \in X}. \tag{8}$$

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**.

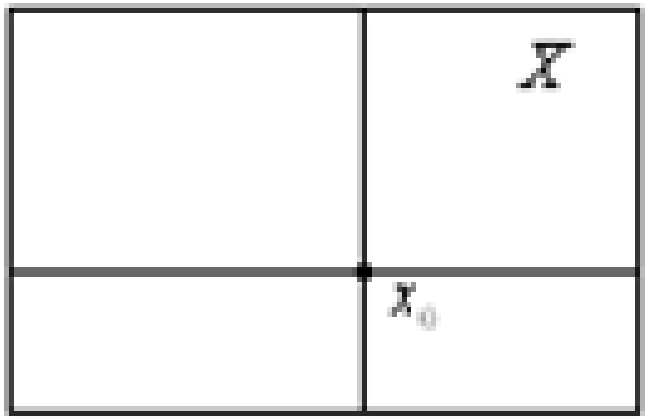
<u>Стохастический метод ветвей и границ.</u>

Пусть имеем задачу без общих ограничений:

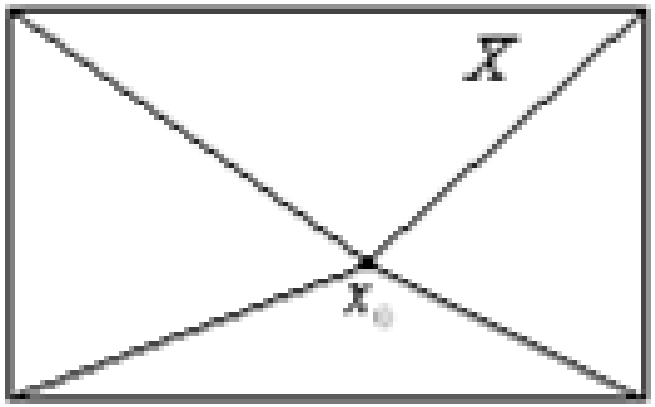
$$F(x) \to \min_{x \in X}, \; X = [a,b]^N. \qquad (9)$$

Способы начального разбиения множества X:

1) <u>Параллелепипедами;</u>



2) <u>Симплексами.</u>

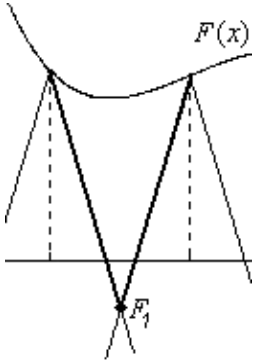


Пусть $\{\varphi(x,y)\}_{y \in X}$ - семейство касательных в точках $y \in X$ минорант для $F(x)$, $\{y \in Z \subseteq X\}$ – конечное или бесконечное множество точек из X.

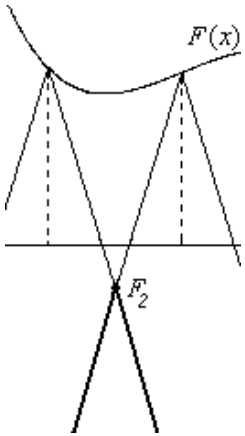
Стохастический метод ветвей и границ.

Минорантные границы оптимальных значений целевой функции на фрагменте разбиения.

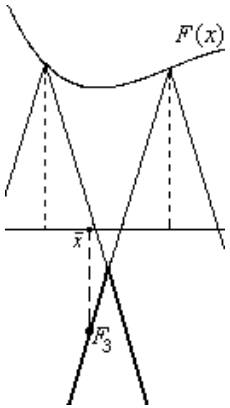
Рассмотрим некоторые оценки снизу для функции $F(x)$:



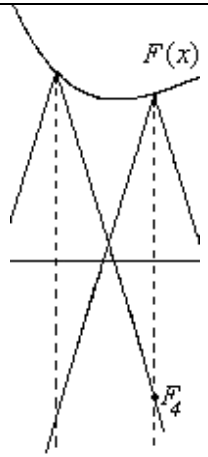
$F_1 = \min_{x \in X} \max_{y \in Z} \varphi(x, y)$



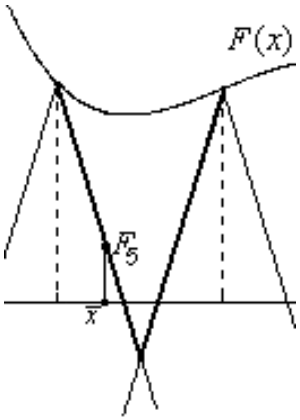
$F_2 = \max_{x \in X} \min_{y \in Z} \varphi(x, y)$



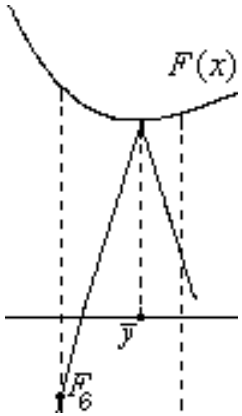
$F_3 = \min_{y \in Z} \varphi(\overline{x}, y)$



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



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



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

<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: ребро наибольшей длины симплекса делятся пополам, и он разбивается на два новых симплекса;

<u>Метод ветвей и приближенных границ.</u>

Рассматривается задача (2) глобальной минимизации непрерывной функции $F(x)$ на множестве $X \subset R^n$:

$$\min_{x \in X} [F(x) = E_x f(x, \theta_x)]$$

Предположения.

а) Пусть $L(Z)$ – точная оценка снизу оптимального значения $F(x)$ на Z, $L(Z) \le \min_{x \in Z} F(x)$, причем $L(Z) = F(Z)$, если множество Z является точкой.

б) Для любой последовательности множеств $X_N \subset X$, $N = 1, 2, ...$, сходящейся в топологии Куратовского к некоторому множеству $Z \in X$ и любой последовательности индексов $M_N(X_N) \to \infty$, имеет место

$$\lim_{N \to \infty} L_{M_N(X_N)}(X_N) \le \lim_{N \to \infty} L(X_N) = L(Z),$$

и, если множество Z является точкой, то

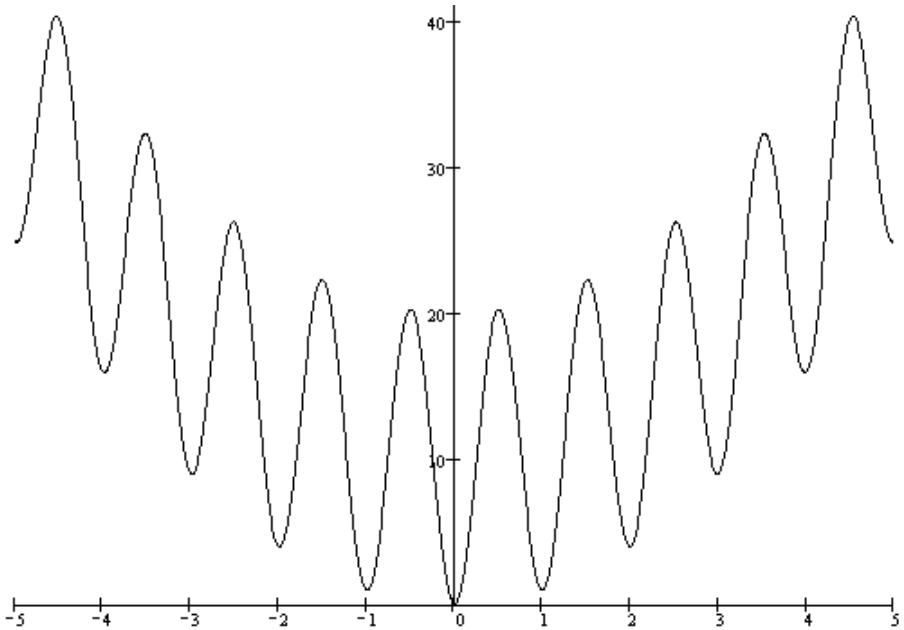
$$\lim_{N \to \infty} L_{M_N(X_N)}(X_N) = \lim_{N \to \infty} L(X_N) = L(Z).$$

Теорема 1 (о сходимости). Пусть $F(x)$ – непрерывная функция, функции множеств $L_N(Z)$ и $L(Z)$ удовлетворяют предположениям а), б). Пусть для любого подмножества разбиения Z, возникающего в ходе работы алгоритма, выполнено $\lim_{N \to \infty} M_N(Z) = \infty$. Предположим, что разбиения осуществляются так, что диаметры получающихся рекордных множеств Y_N стремятся к нулю, при неограниченном числе делений. Тогда последовательность $\{Y_N\}$ бесконечна и для любой предельной точки y рекордных множеств Y_N выполнено: $F(y) = \min_{x \in X} F(x)$.

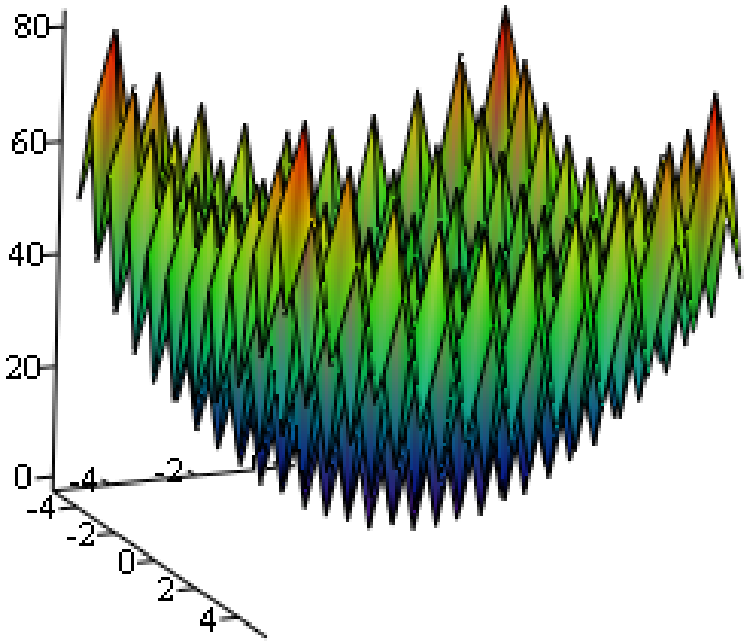
Численные эксперименты.

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, 2	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
	2		Итерации	2555	4531	517	2847
			Оставленные симплексы	27	55	8	49

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

<u>Численные эксперименты.</u>

Стохастическая глобальная оптимизация

(задача размещения сервисных центров).

Предположим, что потребители некоторой услуги распределены в области $\Omega \subset R^m$ согласно распределению $P(d\omega)$. Пусть стоимость обслуживания клиента, живущего в точке $\omega \in \Omega$ из сервисного центра, расположенного в точке $x_i \in X \subset R^m$, задается функцией $c(x_i, \omega)$, например, $c(x_i, \omega) = c\left(\left\| x_i - \omega \right\|\right)$, в частности,

$$c(x_i, \omega) = \frac{\left\| x_i - \omega \right\|^{\alpha}}{\gamma + \left\| x_i - \omega \right\|^{\beta}}, \ \alpha \geq \beta > 0, \quad \gamma > 0 .$$

Если имеется n сервисных центров в точках $x_1, ..., x_n \in X$, а клиент выбирает ближайший центр, то стоимость обслуживания клиента ω задается функцией

$$f(x, \omega) = \min_{1 \leq i \leq n} c(x_i, \omega), \ x = \left(x_1, ..., x_n\right).$$

Задача состоит в размещении n сервисных центров так, чтобы совокупные ожидаемые затраты обслуживания всех потребителей услуги были минимальны:

$$F(x) = \int_{\Omega} \left\{ \min_{1 \leq i \leq n} c(x_i, \omega) \right\} P\left(d\omega\right) \to \min_{x \in D},$$

где D – множество допустимых позиций центров. Координаты центров могут удовлетворять некоторым дополнительным ограничениям, например, они могут быть упорядочены по отношению к некоторому направлению $d \in R^m$, тогда

$$D = \left\{ x = \left(x_1, ..., x_n\right) \mid x_1 \in X, ..., x_n \in X; dx_i \leq dx_{i+1}, i = 1, ..., n-1 \right\}.$$

<u>Численные эксперименты.</u>

Стохастическая глобальная оптимизация

(задача размещения сервисных центров).

В задаче $\alpha = \beta = 2$, а $\gamma = 0.1$. Размещение потребителей $\omega_1, \ldots, \omega_m$ – равномерно распределенные случайные величины на отрезке [0,1] с вероятностями $p_{\omega_1}, \ldots p_{\omega_m}$ соответственно. Для $m = 20$ значения ω_i и p_{ω_i} представлены в таблице 1.

ТАБЛИЦА 1. Распределение потребителей

Значение	0.037991	0.100437	0.193362	0.260176	0.326991
Вероятность	2.15658E-02	3.55347E-03	8.26707E-02	9.97986E-02	2.58563E-02
Значение	0.442882	0.453538	0.480874	0.518605	0.518865
Вероятность	1.20332E-01	1.42203E-01	9.61612E-02	7.31521E-02	2.01054E-02
Значение	0.595109	0.606194	0.646899	0.720796	0.747336
Вероятность	7.57130E-02	8.86943E-02	4.32473E-02	1.66506E-02	3.97192E-02
Значение	0.835398	0.847773	0.885398	0.933185	0.999739
Вероятность	2.55506E-02	1.52854E-02	8.41436E-03	1.32620E-03	7.08675E-08

ТАБЛИЦА 2. Результаты работы алгоритма

n	Вид минорант	Итерации	Оставленные симплексы	X^*
1	1	2091	1301	(0.485959)
	2	18	5	(0.486084)
	3	19	4	(0.486084)
2	1	349550	261436	(0.227470; 0.528999)
	2	329	23	(0.227783; 0.529541)
	3	865	111	(0.228043; 0.529785)
3	2	9289	577	(0.227539; 0.469727; 0.657715)
	3	13592	1128	(0.227417; 0.469604; 0.657654)
4	2	588145	34777	(0.227539; 0.469727; 0.610352; 0.790527)
	3	668906	13106	(0.227661; 0.469604; 0.610229; 0.790649)

<u>Численные эксперименты.</u>

Стохастическая глобальная оптимизация

(максимизация среднего времени жизни сети).

Задача имеет вид:

$$\max_{x\in X}\left[F(x)=E_{\tau}f(\tau)\right],$$

где $f(\tau)$ – функция случайного времени жизни сети; x – вектор оптимизируемых параметров, например, инвестиции ресурсов в элементы сети; E_{τ} – символ математического ожидания по τ. Функция $f(\tau)$ имеет вид:

$$f(\tau)=\max_{z\in Z}\min_{e\in z}\tau_{ze}(x,\omega),$$

где z – путь из конечного множества путей Z, соединяющих "вход" и "выход" сети; $\tau_{ze}(x,\omega)=t(x_{ze})\ln\dfrac{1}{1-\omega_{ze}}$ – случайное экспоненциально распределенное со средним $t(x_{ze})=c_{ze}x_{ze}$ время жизни элемента e пути z; $\omega=\{\omega_{ze}\}$ – вектор независимых случайных чисел, равномерно распределенных на [0,1]; c_{ze} – неотрицательные действительные числа. Очевидно, случайное время жизни элемента сети зависит как от случайного параметра ω_{ze}, так и от вложенных в него инвестиций x_{ze}. Проблема состоит в нахождении глобального максимума $F(x)$ на допустимом множестве X. При этом

$$X=\left\{x\in R^{m\times n}\,\middle|\,x_{ze}\geq 0,\ z=\overline{1,m},\ e=\overline{1,n(z)},\ \sum_{z=1}^{m}\sum_{e=1}^{n(z)}x_{ze}\leq R,\ R\geq 0\right\},$$

где m – количество путей соединяющих "вход" и "выход" сети; $n(z)$ – количество элементов в пути z; R – инвестиции, вложенные в элементы сети.

<u>Численные эксперименты.</u>

Стохастическая глобальная оптимизация

(максимизация среднего времени жизни сети).

ТАБЛИЦА 3. Виды и параметры сетей

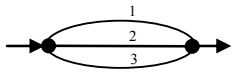
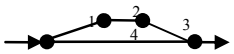
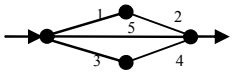
Вид сети	a) 	b) 	c) 
Параметры	c_1=1, c_2=2, c_3=10	c_1=7, c_2=8, c_3=5, c_4=3	c_1=2, c_2=1, c_3=5, c_4=8, c_5=1

ТАБЛИЦА 4. Результаты работы алгоритма

Вид сети	Количество переменных	Количество итераций	x^*
a)	3	325	(0,015254; 0,000218; 0,967967)
b)	4	605	(0,001933; 0,006233; 0,059384; 0,930717)
c)	5	1060	(0,007926; 0,005440; 0,595933; 0,373085; 0,010956)