www.esa-conference.ru

УДК 517.51, УДК 004.94,УДК 519. 6

# Поиск наилучшего приближения в метрике квадратичного отклонения ступенчатыми функциями для распределения Коши

Пастухов Ю.Ф. [1], Пастухов Д.Ф. [1], Чернов С.В. [2], Пастухов А.Ю.
Полоцкий государственный университет, факультет информационных технологий[1]
ОАО «Конструкторское бюро «Дисплей», Витебск [2]

*Научные руководители:* Пастухов Ю.Ф., к. ф. м. н., доцент; Пастухов Д.Ф., к. ф. м. н., доцент, факультет информационных технологий Полоцкого государственного университета

*Магистранты:* Синица П.Р.[1], Субботин А. В.[1], Кохановский А.В.[1], Исаков И.В.[1], Епанешников А.В.[1], Сивограков А.А. [1], Карабанов Р.Ю.[1], Меницкий Е.А. [1], Андреев И.С. [1], Андрейчиков А.Г. [1], Денисова Е.[1], Иваненко Е.С.[1], Карнилович А.В.[1], Петюкевич В. В.[1], Смоляк А.И.[1], Соловьёв А.А.[1], Шевцов М..Ю.[1], Станкевич К.В.[1], Мисевич И.В.[1]

*Предложен метод нахождения наилучшего приближения плотности распределения Коши в пространстве ступенчатых функций на заданном интервале. В данной работе описан метод и алгоритм, заменяющий функцию распределения Коши ступенчатой функцией, являющейся наилучшим приближением плотности распределения Коши в метрике квадратичного отклонения. По сути получен алгоритм квантования функции плотности распределения Коши в пространстве ступенчатых функций на заданном интервале. Данный метод и алгоритм, отличается от алгоритма квантования Ллойда.*

***Ключевые слова:*** *наилучшим приближением функции в метрике квадратичного отклонения, численная аппроксимация интегралов с двенадцатым порядком погрешности, алгоритм Ллойда.*

## 1. Введение

Новым в данной работе является алгоритм нахождения наилучшего приближения плотности распределения Коши в пространстве ступенчатых функций на заданном интервале.

## 2. Квантование функции плотности нормального распределения в метрике квадратичного отклонения

*Определение.* Пусть $m \in N$. Функция $f_m:[a,b] \to R(a<b)$ называется $m$-кусочно-постоянной на [a,b], если $\exists\, x_1 < x_2 < ... < x_{m-1}$ такие что:

$$x_0 = a < x_1 < x_2 < ... < x_{m-1} < b = x_m ,$$

$$f_m(x) = y_i = const \ \forall x \in (x_{i-1}, x_i),\ f_m(x_i) = y_i,\ f_m(x_{i+1}) = y_{i+1}\ y_i \ne y_{i+1},\ \ \forall i = \overline{1, m-1}.$$

Множество $m-$ ступенчатых функций ($m-$ уровней) $f_m:[a,b] \to R(a<b)$ обозначим как $S_m[a,b]$.

Пусть $f:[a,b] \to \Re$, $f \in C^2[a,b]$, $f'(x) < 0\ \forall x \in [a,b]$, $m \in N$. Для минимизации ошибки квантования требуется в пространстве $m-$ступенчатых функций найти наилучшее приближение $h_m:[a,b] \to R$ функции $f:[a,b] \to \Re$ в метрике квадратичного отклонения, такое что $dist = \left\| f - h_m \right\|_{C^2_{[a,b]}} = \min\limits_{f_m \in S_m[a,b]} \left\| f - f_m \right\|_{C^2_{[a,b]}}$. С учетом этого, расстояние оценивается как:

$$dist = \left\| f - h_m \right\|_{C^2_{[a,b]}} = \sqrt{\int_a^b \left(f(x) - h_m(x)\right)^2 dx} = \min_{f_m \in S_m[a,b]} \sqrt{\int_a^b \left(f(x) - f_m(x)\right)^2 dx} = \min_{f_m \in S_m[a,b]} \left\| f - f_m \right\|_{C^2_{[a,b]}}$$

Пусть ступенчатая функция $h_m(x) = y_k$ равна константе на отрезке $x \in (x_{k-1}, x_k)$, $k = \overline{1,m}$, при этом функция ошибки $G(x_1,..., x_{m-1,} y_1,..., y_m) = \sum\limits_{k=1}^{m} \int\limits_{x_{k-1}}^{x_k} \left(f(x) - y_k\right)^2 dx$ описывает квадрат отклонения ступенчатой функции $h_m:[a,b] \to R$ от функции нормального распределения $f:[a,b] \to \Re$. Необходимое условие экстремума функции $G(x_1,..., x_{m-1,} y_1,..., y_m)$описывается системой уравнений:

$$\frac{\partial G(x_1,..., x_{m-1,} y_1,..., y_m)}{\partial x_i} \equiv G^{'}_{x_i} = 0, i = \overline{1, m-1},\ G^{'}_{y_i} = 0,\ \ i = \overline{1,m},$$

Получим явный вид этих уравнений.

Пусть $H(x) = \int\limits_{g_1(x)}^{g_2(x)} f(y,x) dy$ Известно, что

$$\frac{dH(x)}{dx} = -f(g_1(x),x)\frac{dg_1(x)}{dx} + f(g_2(x),x)\frac{dg_2(x)}{dx} + \int_{g_1(x)}^{g_2(x)} \frac{\partial f(y,x)}{\partial x}dy$$

Пусть $H_k(x_{k-1},x_k,y_k) = \int_{x_{k-1}}^{x_k}(f(x)-y_k)^2 dx$ Так как $\frac{dx_{k-1}}{dy_k} = \frac{dx_k}{dy_k} = 0$ , то

$$\frac{\partial H_k(x_{k-1},x_k,y_k)}{\partial y_k} = -(f(x_{k-1})-y_k)^2\frac{dx_{k-1}}{dy_k} + (f(x_k)-y_k)^2\frac{dx_k}{dy_k} + \int_{x_{k-1}}^{x_k}\frac{\partial}{\partial y_k}(f(x)-y_k)^2 dx =$$

$$= \int_{x_{k-1}}^{x_k}\frac{\partial}{\partial y_k}(f(x)-y_k)^2 dx = (-2)\int_{x_{k-1}}^{x_k}(f(x)-y_k)dx = 0 \Rightarrow \int_{x_{k-1}}^{x_k} f(x)dx = \int_{x_{k-1}}^{x_k} y_k dx = y_k(x_k - x_{k-1})$$

$$\frac{\partial H_k(x_{k-1},x_k,y_k)}{\partial x_{k-1}} = -(f(x_{k-1})-y_k)^2\frac{dx_{k-1}}{dx_{k-1}} + (f(x_k)-y_k)^2\frac{dx_k}{dx_{k-1}} + \int_{x_{k-1}}^{x_k}\frac{\partial}{\partial x_{k-1}}(f(x)-y_k)^2 dx = -(f(x_{k-1})-y_k)^2$$

При подстановке вместо $k$ $k+1$ получим

$$\frac{\partial H_{k+1}(x_{k+1-1},x_{k+1},y_{k+1})}{\partial x_{k+1-1}} = \frac{\partial H_{k+1}(x_k,x_{k+1},y_{k+1})}{\partial x_k} = -(f(x_{k+1-1})-y_{k+1})^2 = -(f(x_k)-y_{k+1})^2$$

$$\frac{\partial H_k(x_{k-1},x_k,y_k)}{\partial x_k} = -(f(x_{k-1})-y_k)^2\frac{dx_{k-1}}{dx_k} + (f(x_k)-y_k)^2\frac{dx_k}{dx_k} + \int_{x_{k-1}}^{x_k}\frac{\partial}{\partial x_k}(f(x)-y_k)^2 dx = (f(x_k)-y_k)^2$$

$$G(x_1,\dots,x_{m-1},y_1,\dots,y_m) = \sum_{k=1}^{m}\int_{x_{k-1}}^{x_k}(f(x)-y_k)^2 dx = \sum_{k=1}^{n} H_k(x_{k-1},x_k,y_k)$$

$$\frac{\partial G(x_1,\dots,x_{m-1},y_1,\dots,y_m)}{\partial y_k} = \frac{\partial H_k(x_{k-1},x_k,y_k)}{\partial y_k} = (-2)\int_{x_{k-1}}^{x_k}(f(x)-y_k)dx = 0 \Rightarrow \int_{x_{k-1}}^{x_k} y_k dx = y_k(x_k - x_{k-1})$$

$$\frac{\partial G(x_1,\dots,x_{m-1},y_1,\dots,y_m)}{\partial x_k} = \frac{\partial H_k(x_{k-1},x_k,y_k)}{\partial x_k} - \frac{\partial H_{k+1}(x_k,x_{k+1},y_{k+1})}{\partial x_k} = (f(x_k)-y_k)^2 - (f(x_k)-y_{k+1})^2 =$$

$$= (f(x_k)-y_k+f(x_k)-y_{k+1})(f(x_k)-y_k-f(x_k)+y_{k+1}) = (2f(x_k)-y_k-y_{k+1})(y_{k+1}-y_k) = 0$$

Так как $y_{k+1} \neq y_k \Rightarrow 2f(x_k)-y_k-y_{k+1} = 0 \Rightarrow f(x_k) = \frac{1}{2}(y_k+y_{k+1})$

Отсюда следует следует:

$$\begin{cases} f(B_i) = \frac{1}{2}(C_i + C_{i+1}),\ i = \overline{1,n-1} \\ f(B_i) = \frac{1}{2}C_n \\ \int_{B_{j-1}}^{B_j} f(x)dx = C_j(B_j - B_{j-1}),\ j = \overline{1,n} \end{cases} \qquad (2)$$

Для n+1 ненулевой ступеньки система (1) имеет вид :

$$\begin{cases} f(B_i) = \frac{1}{2}(C_i + C_{i+1}),\ i = \overline{1,n} \\ \int_{B_{j-1}}^{B_j} f(x)dx = C_j(B_j - B_{j-1}),\ j = \overline{1,n+1} \end{cases} \qquad (3)$$

И содержит 2n+1 уравнений и 2n+1 неизвестных.

В работе при вычислении интегралов был использован алгоритм для составной интегральной квадратурной формулы с 12 порядком погрешности, когда исходный отрезок интегрирования делится на число частей кратное десяти (11 узлов равномерной сетки на каждой части). $C_i, x_i, r$ - соответственно веса, узлы и невязка квадратурной формулы.

$$\int_a^b f(z)dz = \sum_{i=0}^{n} C_i f(x_i) + r(f) \quad (4)$$

Интегрируя степенные координатные функции $z^{2s}$ на каноническом отрезке[-1,1], $n_0 = 10$ число частей, на которое делится отрезок[-1,1], учитывая симметрию весов относительно центрального узла $z = 0$ получим:

$$\begin{cases} \int_{-1}^{1} dz = 2 = C_0 + 2\sum_{k=1}^{n_0/2} C_k \\ \int_{-1}^{1} z^{2s} dz = 2/(2s+1) = 2\sum_{k=1}^{n_0/2} C_k \left(2k/n_0\right)^{2s}, s = \overline{1, n_0/2} \end{cases} \quad (5)$$

Для канонического отрезка [-1,1] запишем квадратурную формулу в виде эквивалентном (4)

$$\int_{-1}^{1} f(z)dz = \frac{hn_0}{2}\sum_{i=0}^{n_0} C_i f(x_i) = 5h\sum_{i=0}^{10} C_i f(x_i), \quad \frac{hn_0}{2} = 1, x_i = -1 + ih, i = \overline{0, n_0} \quad (6)$$

Где h - шаг интегрирования, $n_0 = 10$ число отрезков, на которое делится канонический отрезок [-1,1] и каждая часть из k в составной формуле исходного отрезка [a,b] .

А определённый интеграл на отрезке$[a,b]$ отличается от(6) на отрезке [-1,1]длиной интервала в $k = n/n_0 = \frac{b-a}{2}$ раз, используем замену переменных и формулу(4)

$$x = \frac{b+a}{2} + \left(\frac{b-a}{2}\right) z, a \le x \le b, -1 \le z \le 1, dx = \left(\frac{b-a}{2}\right) dz = kdz$$

$$\int_a^b f(x)dx = \int_{-1}^{1} f(z)\left(\frac{b-a}{2}\right)dz = \frac{hn_0}{2}\sum_{i=0}^{n_0 * k} C_i f(x_i), \sum_{i=0}^{n_0} C_i = 2, h = \frac{b-a}{n}, hn = b-a, x_i = a + ih, i = \overline{0, n}$$

Разбивая канонический отрезок[-1,1] $n_0 = 10$ равных частей (из соображений удобства разбиения), используя симметрию весовых коэффициентов, можно получить решение системы уравнений(5) ($n_0 = 10$), в которой 6 неизвестных коэффициентов $C_0$, $C_1$, $C_2$,$C_3$ $C_4$, $C_5$ являются решением системы $n_0/2 + 1 = 6$ линейных неоднородных уравнений с 11 алгебраическим порядком точности:

$$\left\{ C_0 = \frac{17807}{12474}, C_1 = -\frac{4825}{5544}, C_2 = \frac{5675}{6237}, C_3 = -\frac{16175}{99792}, C_4 = \frac{26575}{74844}, C_5 = \frac{16067}{299376} \right. \quad (7)$$

Проверим на компьютере, что рациональный вид коэффициентов (7)(символьное решение системы (5) для $n_0 = 10$) удовлетворяет с двойной точностью(16 значащих цифр). В таблице 1 в левой части указано точное значение интеграла $a(s) = \int_{-1}^{1} z^s dz, s = \overline{0,12}$, а справа численное значение правой части уравнений системы (5) - $b(s)$ с использованием значений весовых коэффициентов (7) ($s -$ показатель степенной функции).

**Таблица 1.** Сравнение интеграла от координатной степенной функции и квадратурной интегральной формулы

| | |
|---|---|
| a(0)=2.0000000000000000 | b(0)=2.0000000000000004 |
| a(1)=0.0000000000000000 | b(1)=0.0000000000000000 |
| a(2)=0.6666666666666666 | b(2)=0.6666666666666669 |
| a(3)=0.0000000000000000 | b(3)=-0.0000000000000000 |
| a(4)=0.4000000000000000 | b(4)=0.4000000000000001 |
| a(5)=0.0000000000000000 | b(5)=-0.0000000000000000 |
| a(6)=0.2857142857142857 | b(6)=0.2857142857142858 |
| a(7)=0.0000000000000000 | b(7)=0.0000000000000000 |
| a(8)=0.2222222222222222 | b(8)=0.2222222222222223 |
| a(9)=0.0000000000000000 | b(9)=-0.0000000000000000 |
| a(10)=0.1818181818181818 | b(10)=0.1818181818181819 |
| a(11)= .0000000000000000 | b(11)=-.0000000000000000 |
| a(12)= .1538461538461539 | b(12)= .1554621683809524 |

Из таблицы 1 видно, что алгебраический порядок точности системы уравнений (5) при $n_0 = 10$ равен11, а порядок погрешности квадратурной формулы $\int_{-1}^{1} f(z)dz = 5h\sum_{i=0}^{10} C_i f(x_i), 5h = 1, \sum_{i=0}^{10} C_i = 2, x_i = -1 + ih, i = \overline{0,10}$

равен 12( $C_i$ определяются с помощью (7)).

Из (6) для $n_0 = 10$ получим составную формулу:

$$\int_a^b f(z)dz = 5h\sum_{i=0}^{n} C_i f(x_i), h = \frac{(b-a)}{n}, \sum_{i=0}^{10} C_i = 2, x_i = a + ih, n = 10k \quad (8)$$

в которой весовые коэффициенты $C_i$ определяются алгоритмом (9):

$$\begin{cases} \text{если } \ j = 0 \ \text{ или } j = n : C_j = \dfrac{16067}{299376}; \\ \text{если } \ j \equiv 1 \bmod 10 \ \text{ или } \ j \equiv 9 \bmod 10 : C_j = \dfrac{26575}{74844}; \\ \text{если } \ j \equiv 2 \bmod 10 \ \text{ или } \ j \equiv 8 \bmod 10 : C_j = -\dfrac{16175}{99792}; \\ \text{если } \ j \equiv 3 \bmod 10 \ \text{ или } \ j \equiv 7 \bmod 10 : C_j = \dfrac{5675}{6237}; \\ \text{если } \ j \equiv 4 \bmod 10 \ \text{ или } \ j \equiv 6 \bmod 10 : C_j = -\dfrac{4825}{5544}; \\ \text{если } j \equiv 5 \bmod 10 : C_j = \dfrac{17807}{12474}; \\ \text{если } \ j \equiv 0 \bmod 10 \ , j > 0, j < n : C_j = \dfrac{16067}{149688}; \end{cases} \quad (9)$$

Методы точных вычислений в стеганогафии описаны также в работах[6-9].

Для исследования рассмотренного алгоритма разработана программа, которая позволяет получать пороговые уровни и ошибку приближения для различного числа n (таблица 2) с учетом интегральных квадратурных формул(8, 9).

**Таблица 2.** Результаты исследований ошибки квантования распределения Коши для различного количества уровней

| *n* | 8 | 16 | 32 | 64 | 128 |
|---|---|---|---|---|---|
| *G= dist*$^2$ | $2.26*10^{-3}$ | $6,14 * 10^{-4}$ | $1,61* 10^{-4}$ | $4,51* 10^{-5}$ | $1,41* 10^{-5}$ |
| *dist* | $4,76* 10^{-2}$ | $2,48* 10^{-2}$ | $1,27* 10^{-2}$ | $6,72* 10^{-3}$ | $3,75* 10^{-3}$ |

На рисунке 1 представлен пример квантования функции плотности распределения Коши $f(x) = \dfrac{1}{\pi a\left(1 + \left(\dfrac{x-b}{a}\right)^2\right)} = \dfrac{1}{\pi(1+x^2)}$ *при* $a = 1$ , $b = 0$ на основе предложенного подхода в метрике квадратичного отклонения для числа ступеней а)m=10 и б)m=20.(на одну сторону). Полученные значения пороговых уровней для n=20(m=10 ступеней): 0; 0.264; 0.387; 0.504; 0.616; 0.729; 0.848; 0.972; 1.113; 1.261; 1.440; 1.632; 1.881; 2.155; 2.543; 2.986; 3.709; 4.588; 6.527; 9.250

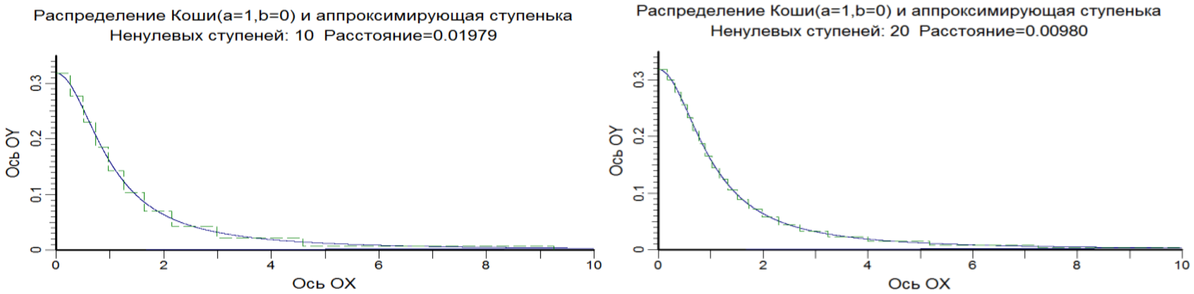
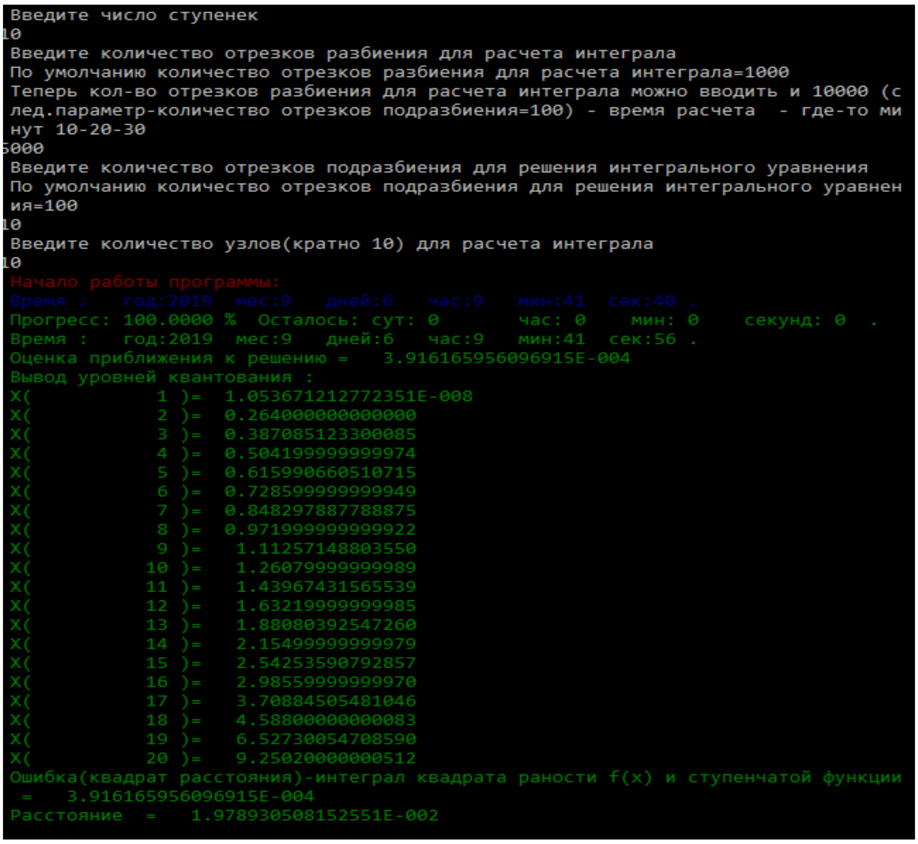



**Рис. 1.** Результат квантования: а)для m=10; б)для m=20

www.esa-conference.ru



```
 Введите число ступенек
10
 Введите количество отрезков разбиения для расчета интеграла
 По умолчанию количество отрезков разбиения для расчета интеграла=1000
 Теперь кол-во отрезков разбиения для расчета интеграла можно вводить и 10000 (с
 лед.параметр-количество отрезков подразбиения=100) - время расчета  - где-то ми
 нут 10-20-30
5000
 Введите количество отрезков подразбиения для решения интегрального уравнения
 По умолчанию количество отрезков подразбиения для решения интегрального уравнен
 ия=100
10
 Введите количество узлов(кратно 10) для расчета интеграла
10
 Начало работы программы:
 Время :   год:2019  мес:9   дней:6   час:9   мин:41  сек:40 .
 Прогресс: 100.0000 %  Осталось: сут: 0        час: 0    мин: 0    секунд: 0  .
 Время :   год:2019  мес:9   дней:6   час:9   мин:41  сек:56 .
 Оценка приближения к решению =   3.916165956096915E-004
 Вывод уровней квантования :
 X(           1 )=  1.053671212772351E-008
 X(           2 )=  0.264000000000000
 X(           3 )=  0.387085123300085
 X(           4 )=  0.504199999999974
 X(           5 )=  0.615990660510715
 X(           6 )=  0.728599999999949
 X(           7 )=  0.848297887788875
 X(           8 )=  0.971999999999922
 X(           9 )=   1.11257148803550
 X(          10 )=   1.26079999999989
 X(          11 )=   1.43967431565539
 X(          12 )=   1.63219999999985
 X(          13 )=   1.88080392547260
 X(          14 )=   2.15499999999979
 X(          15 )=   2.54253590792857
 X(          16 )=   2.98559999999970
 X(          17 )=   3.70884505481046
 X(          18 )=   4.58800000000083
 X(          19 )=   6.52730054708590
 X(          20 )=   9.25020000000512
 Ошибка(квадрат расстояния)-интеграл квадрата раности f(x) и ступенчатой функции
  =   3.916165956096915E-004
 Расстояние  =   1.978930508152551E-002
```

**Рис. 3.** Уровни квантования для m=10

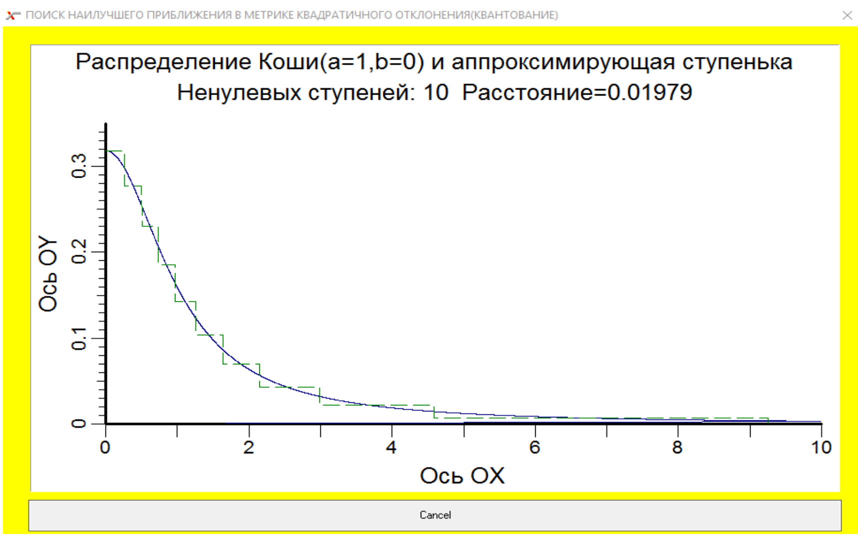



**Рис. 4.** График уровней квантования: для m=10

На рисунке 3,4 приведен результат работы программы для поиска оптимальных уровней квантования для числа ступеней m=10

**Литература:**

1. Пастухов Ю.Ф., Пастухов Д.Ф., Богуш Р.П., Пастухов А.Ю. Определение оптимальных уровней восстановления и квантования плотности нормального распределения в метрике квадратичного отклонения для алгоритма сжатия данных/Пастухов Ю.Ф.//Евразийское Научное Объединение. 2018. Т. 1. № 11(45). С. 16-21.

2. Моделирование сжатия радиолокационных данных дистанционного зондирования Земли на основе блочного адаптивного квантования / Богуш Р. П., [и др.] // Вестник ПГУ. Серия С. Фундаментальные науки. 2019. №4 – С. 7-15.

3. Карабанов Р. Ю., Пастухов Д. Ф., Пастухов Ю. Ф., Богуш Р. П. Наилучшее приближение монотонно убывающих функций кусочно – постоянными функциями в метрике квадратичного отклонения. В сборнике: Прикладная математика и информатика: современные исследования в области естественных и технических наук. Сборник научных статей IV научно- практической международной конференции (школы-семинара) молодых ученых: в двух частях. 2018. С. 48-5

4. Пастухов Ю. Ф. Необходимые условия в обратной вариационной задаче/ Ю.Ф. Пастухов // Фундаментальная и прикладная математика. 7:1 (2001). С. 285–288.

5. Пастухов Ю.Ф., Пастухов Д.Ф., Богуш *Р.П.* Квантование функции плотности нормального распределения в метрике квадратичного отклонения. В сборнике: Информационно-коммуникационные технологии: достижения, проблемы, инновации (ИКТ-2018). Электронный сборник статей I международной научно-практической конференции, посвященной 50-летию Полоцкого государственного университета. 2018. С. 92-95.

6. Пастухов Д.Ф. Оптимальный порядок аппроксимации разностной схемы волнового уравнения на отрезке / Д.Ф. Пастухов, Ю.Ф. Пастухов, Н.К. Волосова // Вестник Полоцкого университета. Серия С. Фундаментальные науки. – 2018. – № 12. – С. 60–74.

7. Пастухов Д.Ф. К вопросу о редукции неоднородной краевой задачи Дирихле для волнового уравнения на отрезке / Д.Ф. Пастухов, Ю.Ф. Пастухов, Н.К. Волосова // Вестник Полоцкого университета. Серия С. Фундаментальные науки. – 2018. – № 4. – С. 167–186.

8. Пастухов Д.Ф. Минимальная разностная схема для уравнения Пуассона на параллелепипеде с шестым порядком погрешности / Д.Ф. Пастухов, Ю.Ф. Пастухов, Н.К. Волосова // Вестник Полоцкого университета. Серия С. Фундаментальные науки. – 2019. – № 4. – С. 154–173.

9. Пастухов Д.Ф., Волосова Н.К., Волосова А.К. Некоторые методы передачи Q-R кода с помощью стеганографии / Д.Ф. Пастухов, Н.К. Волосова, А.К. Волосова // Мир транспорта. 2019. Т. 17. № 3.(82). С. 16–39.

10. Вакуленко С. П., Волосова Н. К., Пастухов Д. Ф. Способы передачи Q-R кода с помощью, стеганографии / С. П. Вакуленко, Н.К. Волосова, Д. Ф. Пастухов // Мир транспорта. 2018. Т. 16. № 5(78). С. 14–25.

11. Пастухов Ю.Ф., Соловьев А.А., Карабанов Р.Ю., Субботин А.В., Пастухов Д.Ф. Роль обратной функции при определении оптимальных пороговых уровней восстановления. Вестник современных исследований. 2019. № 2.13 (29). С. 33-41.