Нижегородский государственный университет им. Н.И.Лобачевского

Национальный исследовательский университет

Институт информационных технологий, математики и механики

Системы принятия решений

Vladimir Grishagin

Характеристические алгоритмы оптимизации

$$\varphi(x) \to \min, x \in Q = [a, b] \qquad (2.1)$$

Определение. Алгоритм решения задачи (*) называется характеристическим, если, начиная с некоторого шага поиска $k_0 \ge 1$, выбор координаты очередного испытания x^{k+1}, $(k \ge k_0)$ заключается в выполнении следующих действий.

1. Задать множество

$$\Lambda_k = \{x_0, x_1, \ldots, x_\tau\}$$

конечного числа $\tau + 1 = \tau(k) + 1$ точек области $Q = [a, b]$, полагая, что

$a \in \Lambda_k, b \in \Lambda_k$, все координаты предшествующих испытаний $x^i \in \Lambda_k, 1 \le i \le k,$

множество Λ_k упорядочено (нижним индексом) по возрастанию координаты, т.е.

$$a = x_0 < x_1 < \ldots < x_{\tau-1} < x_\tau = b.$$

и сопоставить точкам множества значения $z_i = \varphi(x_i), 0 \le i \le \tau.$

Характеристические алгоритмы оптимизации

2. Каждому интервалу (x_{i-1}, x_i), $1 \le i \le \tau$, поставить в соответствие число $R(i)$, называемое характеристикой этого интервала.

3. Определить интервал (x_{t-1}, x_t), которому соответствует максимальная характеристика $R(t)$, т.е.

$$R(t) = \max\{R(i) : 1 \le i \le \tau\}$$

4. Провести очередное испытание в точке

$$x^{k+1} = d(t) \in (x_{t-1}, x_t)$$

и вычислить значение $z^{k+1} = \varphi(x^{k+1})$.

Двусторонняя сходимость.

$$\varphi(x) \to \min, x \in Q = [a,b] \qquad (2.1)$$

Теорема 2.1. Пусть точка x^* является предельной точкой (точкой накопления) последовательности испытаний $\{x^k\}$, генерируемой характеристическим алгоритмом при решении задачи (2.1), причем $x^* \neq a$, $x^* \neq b$. Если характеристики $R(i)$ и правило $d(\cdot)$ выбора координаты очередного испытания x^{k+1} обеспечивают выполнение условий:

i) если точка $\bar{x} \in [x_{i(k)-1}, x_{i(k)}]$ и $x_{i(k)-1} \to \bar{x}$, $x_{i(k)} \to \bar{x}$, когда $k \to \infty$, тогда

$$R(i(k)) \to -\mu\,\varphi(\bar{x}) + c \qquad (2.2)$$

ii) если, начиная с некоторого шага поиска, интервал $(x_{i-1}, x_i), i = i(k)$, не будет содержать точек испытаний, т.е. существует номер $k_s \geq 1$ такой, что для всех $k \geq k_s$

$$(x_{i-1}, x_i) \cap \{x^k\} = \varnothing \qquad (2.3)$$

тогда для характеристики этого интервала выполняется строгое неравенство

$$\lim_{k\to\infty} R(i) > -\mu \min\{\varphi(x_{i-1}), \varphi(x_i)\} + c \qquad (2.4)$$

iii) для точки нового испытания x^{k+1} имеет место неравенство

$$\max\{x^{k+1} - x_{t-1}, x_t - x^{k+1}\} \leq \nu(x_t - x_{t-1}) \qquad (2.5)$$

где в (2.2), (2.4), (2.5) величины μ, c, и ν некоторые константы, причем

$$\mu \geq 0,\ 0 < \nu < 1 \qquad (2.6)$$

тогда последовательность $\{x^k\}$ содержит две подпоследовательности, одна из которых сходится к x^* слева, другая справа.

Двусторонняя сходимость.

Следствие. В вычислительную схему алгоритма, удовлетворяющего условиям Теоремы 2.1, может быть введено условие остановки вида

$$x_t - x_{t-1} \le \varepsilon \qquad (2.7)$$

где t - это номер интервала с максимальной характеристикой, а $\varepsilon > 0$ - заданная точность поиска.

Двусторонняя сходимость (перебор)

i) Если точка $\bar{x} \in [x_{i-1}, x_i]$ и $x_{i-1} \to \bar{x}, \quad x_i \to \bar{x}$, когда $k \to \infty$, тогда

$$R(i) = x_i - x_{i-1} \to 0, \; m.e. \; \mu = 0, \, c = 0$$

ii) Если $(x_{i-1}, x_i) \cap \{x^k\} = \varnothing$, то требуется

$$\lim_{k \to \infty} R(i) > -\mu \min\{z_{i-1}, z_i\} + c = 0$$

iii) для точек новых испытаний

$$\max\{x^{k+1} - x_{t-1}, x_t - x^{k+1}\} \le \nu (x_t - x_{t-1})$$

$$x^{k+1} = 0.5(x_{t-1} + x_t).$$

$$x^{k+1} - x_{t-1} = x_t - x^{k+1} = 0.5(x_t - x_{t-1})$$

$$\nu = 0.5$$

Двусторонняя сходимость (метод Пиявского)

$$R(i) = 0.5m(x_i - x_{i-1}) - (z_i + z_{i-1})/2,$$

$$x^{k+1} = 0.5(x_t + x_{t-1}) - (z_t - z_{t-1})/(2m),$$

Условие Липшица

$$\left|\varphi(x') - \varphi(x'')\right| \le L\left|x' - x''\right|, x', x'' \in [a, b]$$

i) если точка $\bar{x} \in [x_{i-1}, x_i]$ и $x_{i-1} \to \bar{x},\ x_i \to \bar{x}$ когда $n \to \infty$ тогда

$$R(i) \to -\varphi(\bar{x})$$

$$\mu = 1, c = 0$$

Двусторонняя сходимость (метод Пиявского)

ii) если $(x_{i-1}, x_i) \cap \{x^k\} = \varnothing$
тогда

$$\lim_{n\to\infty} R(i) > -\mu \min\{z_{i-1}, z_i\}$$

$$\min\{z_{i-1}, z_i\} = \frac{1}{2}(z_{i-1} + z_i - |z_{i-1} - z_i|)$$

Возьмем $m > L.$ Вследствие условия Липшица

$$L(x_i - x_{i-1}) \geq |z_{i-1} - z_i|$$

$$m\frac{x_i - x_{i-1}}{2} - \frac{z_i + z_{i-1}}{2} > L\frac{x_i - x_{i-1}}{2} - \frac{z_i + z_{i-1}}{2} \geq \frac{|z_{i-1} - z_i| - (z_{i-1} + z_i)}{2} = -\min\{z_{i-1}, z_i\}$$

Двусторонняя сходимость (метод Пиявского)

iii) для точек новых испытаний

$$\max\{x^{k+1} - x_{t-1}, x_t - x^{k+1}\} \le \nu(x_t - x_{t-1})$$

$$x_t - x^{k+1} = \frac{x_t - x_{t-1}}{2} + \frac{z_t - z_{t-1}}{2m} \le \frac{x_t - x_{t-1}}{2} + L\frac{x_t - x_{t-1}}{2m} = \left(1 + \frac{L}{m}\right)\frac{x_t - x_{t-1}}{2}$$

$$\nu = 0,5(1 + L / m)$$

Двусторонняя сходимость (метод Стронгина)

$$R(i) = m(x_i - x_{i-1}) + \frac{(z_i - z_{i-1})^2}{m(x_i - x_{i-1})} - 2(z_i + z_{i-1}),$$

$$x^{k+1} = 0.5(x_t + x_{t-1}) - (z_t - z_{t-1})/(2m),$$

Условие Липшица

$$|\varphi(x') - \varphi(x'')| \le L|x' - x''|, x', x'' \in [a,b]$$

i) если точка $\bar{x} \in [x_{i-1}, x_i]$ и $x_{i-1} \to \bar{x}, x_i \to \bar{x}$, когда $n \to \infty$ тогда

$$R(i) \to -4\varphi(\bar{x})$$

$$\mu = 4, c = 0$$

Двусторонняя сходимость (метод Стронгина)

ii) если $(x_{i-1}, x_i) \cap \{x^k\} = \varnothing$

тогда $\lim\limits_{n\to\infty} R(i) > -\mu\xi_i = -4\min\{z_{i-1}, z_i\}$

$$\min\{z_{i-1}, z_i\} = \frac{1}{2}(z_{i-1} + z_i - |z_{i-1} - z_i|)$$

Возьмем $m > L.$

Случай 1. $z_{i-1} = z_i$

$$R(i) = m(x_i - x_{i-1}) - 2(z_i + z_{i-1}) > -2(z_i + z_{i-1}) = -4\min\{z_{i-1}, z_i\}$$

Случай 2. $z_{i-1} \neq z_i$

$$R(i) = |z_i - z_{i-1}|(\beta + \frac{1}{\beta}) - 2(z_i + z_{i-1})$$

$$\beta = \frac{m(x_i - x_{i-1})}{|z_i - z_{i-1}|} > \frac{m}{L} > 1 \qquad \beta + \frac{1}{\beta} > 2$$

$$R(i) > 2|z_i - z_{i-1}| - 2(z_i + z_{i-1}) = -4\min\{z_i, z_{i-1}\}$$

Двусторонняя сходимость (метод Стронгина)

iii) для точек новых испытаний

$$\max\{x^{k+1}-x_{t-1}, x_t - x^{k+1}\} \le \nu(x_t - x_{t-1})$$

$$x_t - x^{k+1} = \frac{x_t - x_{t-1}}{2} + \frac{z_t - z_{t-1}}{2m} \le \frac{x_t - x_{t-1}}{2} + L\frac{x_t - x_{t-1}}{2m} = \left(1+\frac{L}{m}\right)\frac{x_t - x_{t-1}}{2}$$

$$\nu = 0,5(1+L/m)$$

$$m = \begin{cases} rM, M>0 \\ 1, M=0 \end{cases}$$

$$M = \max_{1\le i\le \tau}\frac{|z_i - z_{i-1}|}{x_i - x_{i-1}}$$

$r>1$ - параметр метода

$$x_t - x^{k+1} = \frac{x_t - x_{t-1}}{2} + \frac{z_t - z_{t-1}}{2m} \le \frac{x_t - x_{t-1}}{2} + M\frac{x_t - x_{t-1}}{2rM} = \left(1+\frac{1}{r}\right)\frac{x_t - x_{t-1}}{2}$$

$$\nu = 0.5(1+1/r)$$

Двусторонняя сходимость (метод Кушнера)

$$R(i) = -\frac{4(\varphi_k^* - \delta_k - z_i)(\varphi_k^* - \delta_k - z_{i-1})}{x_i - x_{i-1}}, 1 \le i \le \tau,$$

$$x^{k+1} = x_{t-1} + \frac{(x_t - x_{t-1})(\varphi_k^* - \delta_k - z_t)}{2(\varphi_k^* - \delta_k) - z_t - z_{t-1}}$$

где $\delta_k > 0$ - параметр метода, а $\varphi_k^* = \min_{0 \le i \le \tau} z_i$.

Потребуем непрерывность целевой функции

i) Пусть точка $\bar{x} \in [x_{i-1}, x_i]$ и $x_{i-1} \to \bar{x}, \ x_i \to \bar{x}$, когда $k \to \infty$,

и, начиная с некоторого шага поиска, параметр метода $\delta_k > \delta > 0$, тогда $\varphi_k^* - z_{i-1} - \delta_k < -\delta < 0$, $\varphi_k^* - z_i - \delta_k < -\delta < 0$ и характеристика

$$R(i(k)) = -\frac{4(\varphi_k^* - \delta_k - z_i)(\varphi_k^* - \delta_k - z_{i-1})}{x_i - x_{i-1}} \to -\infty$$

т.к. длина интервала (x_{i-1}, x_i) стремится к нулю. Следовательно, (2.2) выполняется при $\mu = 0$ и $c = -\infty$ (доказательство теоремы 2.1 не меняется, если предел в правой части (2.2) будет равен −∞).

Двусторонняя сходимость (метод Кушнера)

ii) Если $(x_{i-1}, x_i) \cap \{x^k\} = \varnothing$, то требуется

$$\lim_{k\to\infty} R(i) > -\mu \min\{z_{i-1}, z_i\} + c$$

При $\mu = 0, c = -\infty$ это неравенство имеет вид

$$\lim_{k\to\infty} R(i) > -\infty$$

Но любой интервал, в который испытания больше не попадают, имеет положительную длину, что доказывает требуемое неравенство.

Двусторонняя сходимость (метод Кушнера)

iii) для точек новых испытаний

$$\max\{x^{k+1}-x_{t-1}, x_t - x^{k+1}\} \le \nu (x_t - x_{t-1})$$

Рассмотрим разности

$$x^{k+1}-x_{t-1} = \beta_1 (x_t - x_{t-1}),\;\; x_t - x^{k+1} = \beta_2 (x_t - x_{t-1})$$

где

$$\beta_1 = \frac{z_{t-1}-\varphi_k^* + \delta_k}{z_{t-1}+z_t - 2(\varphi_k^* - \delta_k)},\;\; \beta_2 = \frac{z_t - \varphi_k^* + \delta_k}{z_{t-1}+z_t - 2(\varphi_k^* - \delta_k)}$$

Введем вспомогательную функцию $\sigma(w,\alpha) = \dfrac{w+\alpha}{w+2\alpha}$. Так как $\min\{z_{t-1}, z_t\} \ge \varphi_k^*$, имеем $\beta_1 \le \sigma(z_{t-1}-\varphi_k^*, \delta_k),\, \beta_2 \le \sigma(z_t - \varphi_k^*, \delta_k)$.

Для любого положительного *w* функция $\sigma(w,\alpha)$ убывает по *α*, потому что производная $\sigma'_\alpha(w,\alpha) = \dfrac{-w}{(w+2\alpha)^2} < 0$, откуда для $\delta_k > \delta > 0$

$$\beta_1 \le \sigma(z_{t-1}-\varphi_k^*, \delta),\, \beta_2 \le \sigma(z_t - \varphi_k^*, \delta)$$

Двусторонняя сходимость (метод Кушнера)

Кроме того, функция $\sigma(w,\alpha)$ возрастает по *w* для любого положительного *α* , т.к. производная $\sigma'_w(w,\alpha)=\dfrac{\alpha}{(w+2\alpha)^2}>0$, поэтому

$$\max\{\sigma(z_{t-1}-\varphi_k^*,\delta),\sigma(z_t-\varphi_k^*,\delta)\}\le\sigma(\varphi_{\max}-\varphi_{\min},\delta)$$

Последнее неравенство означает, что

$$\max\{\beta_1,\beta_2\}\le\sigma(\varphi_{\max}-\varphi_{\min},\delta)=\frac{\varphi_{\max}-\varphi_{\min}+\delta}{\varphi_{\max}-\varphi_{\min}+2\delta}$$

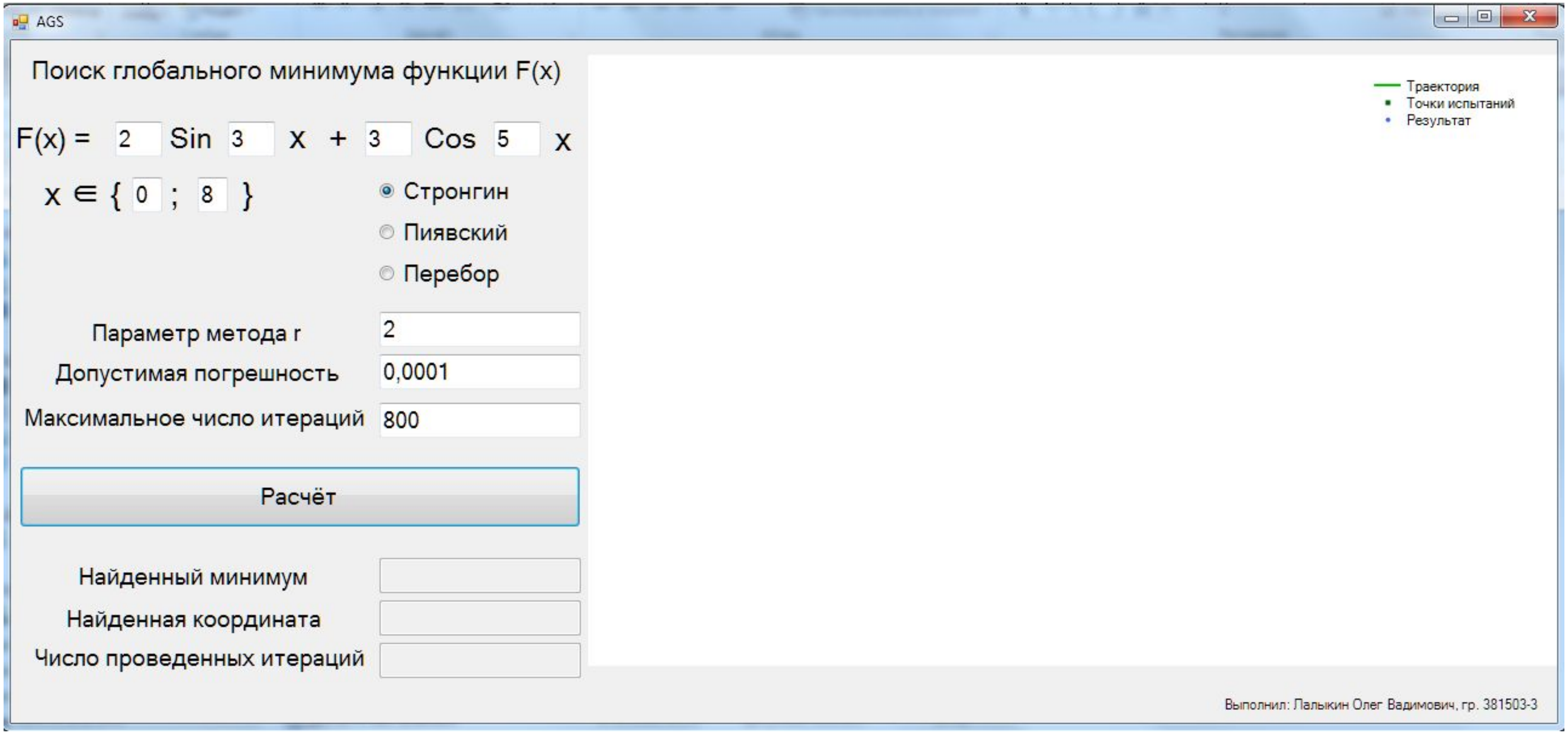
следовательно,

$$x^{k+1}-x_{t-1}\le\frac{\varphi_{\max}-\varphi_{\min}+\delta}{\varphi_{\max}-\varphi_{\min}+2\delta}(x_t-x_{t-1}),\ \ x_t-x^{k+1}\le\frac{\varphi_{\max}-\varphi_{\min}+\delta}{\varphi_{\max}-\varphi_{\min}+2\delta}(x_t-x_{t-1})$$

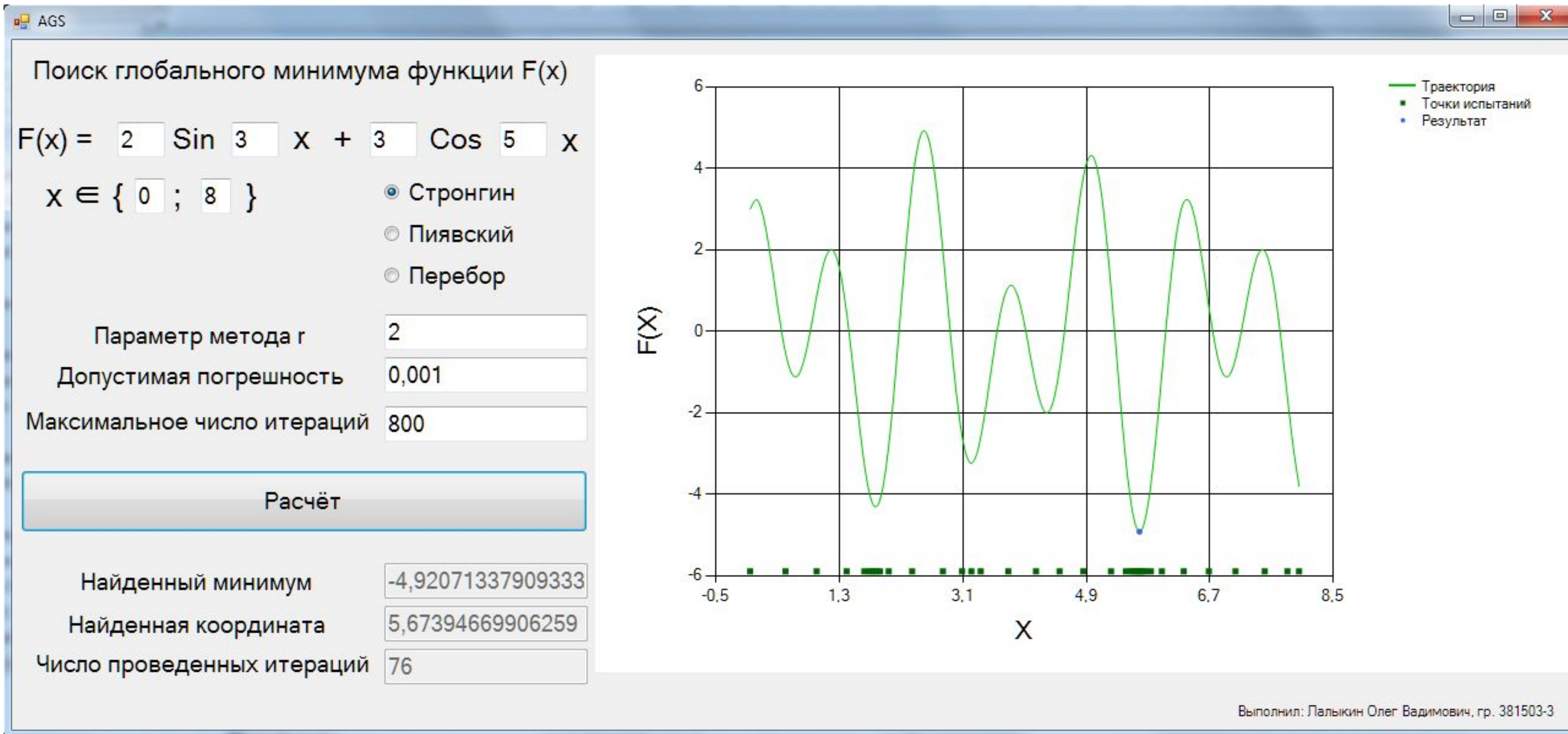
Таким образом, в качестве константы *v* в (2.6) можно взять величину $\dfrac{\varphi_{\max}-\varphi_{\min}+\delta}{\varphi_{\max}-\varphi_{\min}+2\delta}$

которая, очевидно, положительна и меньше единицы.

Программная система



Программная система



$$\varphi(x) = \alpha \sin(\beta x) + \gamma \cos(\delta x)$$