

Нижегородский государственный университет им. Н.И.Лобачевского
Национальный исследовательский университет
Институт информационных технологий, математики и механики

Системы принятия решений

VLADIMIR GRISHAGIN



Оценки экстремума

Рассмотрим следующий простой пример. Пусть класс \mathcal{Q} - класс непрерывных функций, т.е. априорно известно, что минимизируемая функция Q непрерывна в области Ω . Предположим, что вычислены значения функции в конечном числе точек ω_k . Что после этого можно сказать о координате глобального минимума? Каковы бы ни были k точки и значения $y^* \in Q$ (y^* для любой точки ω_k всегда можно построить непрерывную функцию (проходящую через точки ω_k т.е. $\Phi(\omega_k)$ принадлежащую \mathcal{Q} классу из (1.11), которая имеет глобальный минимум в точке ω_k с любым наперед заданным значением ϵ $\min_{1 \leq i \leq k} z^i$).

Например, в качестве такой функции можно взять интерполяционный полином k -й степени, проходящий через точки (ω_k, y^*) , $1 \leq i \leq k$, (y^*, ω_k) .

Оценки экстремума

Все сказанное означает, что по результатам конечного числа испытаний никаких выводов о расположении координаты глобального минимума сделать нельзя. Точно так же о величине глобального минимума можно лишь сказать, что

$$\varphi^* \leq \varphi_k^* \quad (1.17)$$

где φ_k^* из (1.15), однако оценить величину

$$\varepsilon_k = |\varphi^* - \varphi_k^*| \quad (1.18)$$

т.е. погрешность решения задачи, невозможно.

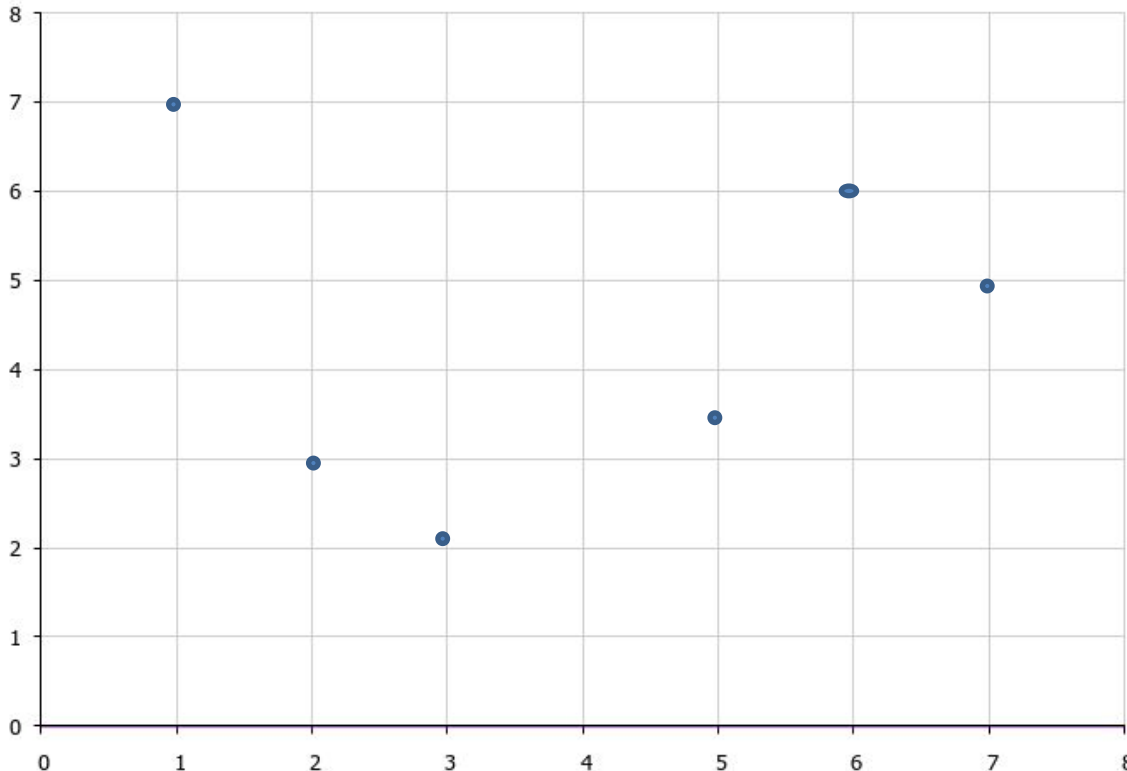
Возможность получения оценок экстремума по конечному числу испытаний зависит от свойств класса функций, которому принадлежит минимизируемая функция, или, другими словами, от априорной информации о функции. Фактически известны лишь два широких класса функций, допускающих построение таких оценок: класс одномерных унимодальных функций и класс функций, удовлетворяющих условию Липшица (в общем случае многомерных и многоэкстремальных).

Оценки экстремума для унимодальных функций

Функция $\varphi(y)$ является строго унимодальной на отрезке $[a, b]$, если существует точка $y^* \in [a, b]$ такая, что для всех $y^* < y' < y'' \leq b$ выполняется $\varphi(y^*) < \varphi(y') < \varphi(y'')$.

а для всех $y^* < y' < y'' \leq b$ справедливо неравенство $\varphi(y^*) < \varphi(y') < \varphi(y'')$.

Пример. Унимодальная функция на отрезке $[0, 8]$.



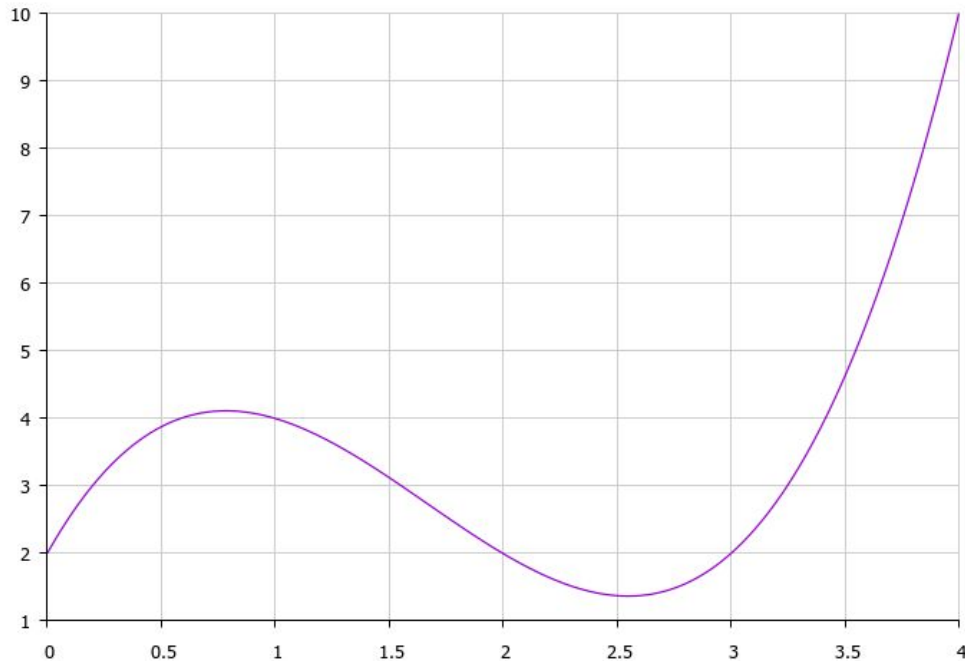
Построение миноранты

Задание

Построить для константы Липшица $L = 15$ миноранту функции $y^3 - 5y^2 + 6y + 2$ на интервале $[0,4]$ по точкам испытаний

$$y_1 = 0, y_2 = 1, y_3 = 2, y_4 = 4$$

и указать оценку величины глобального минимума $\varepsilon_k = \varphi_k^* - \Psi_k^*$



Оптимальность алгоритмов оптимизации

Рассмотрим класс алгоритмов S , предназначенных для решения задач оптимизации функций $V(\varphi, s)$.

Вводится вещественная функция $V(\varphi, s)$ - критерий эффективности решения задачи минимизации функции $V(\varphi, s)$ с помощью метода s .

Критерии эффективности метода s на классе

$$W(s) = \sup_{\varphi \in \Phi} V(\varphi, s)$$

или

$$W(s) = \int V(\varphi, s) dP(\varphi),$$

$P(\varphi)$ — некоторое распределение вероятностей, заданных на классе измеримых подмножеств множества Φ .

Оптимальный алгоритм s^* :

$$W(s^*) = \inf_{s \in S} W(s) \tag{1.28}$$

Оптимальность алгоритмов оптимизации

ε -оптимальный алгоритм оптимизации W_ε^* ($\varepsilon > 0$) :

$$W(s_\varepsilon^*) \leq \inf_{s \in S} W(s) + \varepsilon \quad (1.29)$$

Наилучший (последовательно-оптимальный алгоритм) оптимизации (А.Г.Сухарев) - алгоритм, наилучшим образом использующий апостериорную (поисковую) информацию.

Одношаговая оптимальность

Существуют значительные трудности в создании алгоритмов в соответствии с принципами оптимальности (1.28) или (1.29). Чтобы преодолеть эти трудности, можно попытаться упростить понятие оптимальности. Такую возможность предоставляет принцип одношаговой оптимальности, когда мы требуем оптимального поведения алгоритма не в течение всего процесса оптимизации, а только на один шаг вперед, т.е. при выборе точки очередного испытания.

Чтобы кратко описать модель одношагово-оптимального алгоритма, вспомним ряд обозначений, принятых в общей схеме численного метода оптимизации (1.9)–(1.14). После k испытаний, выполненных методом в точках y^1, y^2, \dots, y^k с результатами $z^1, z^2, \dots, z^k = \varphi(y^i), 1 \leq i \leq k$,

мы имеем в своем распоряжении новую информацию $\omega_k = \{(y^1, z^1), (y^2, z^2), \dots, (y^k, z^k)\}$

и мы можем утверждать, что целевая функция принадлежит классу

$$\Phi(\omega_k) = \{\psi \in \Phi : \psi(y^i) = z^i, 1 \leq i \leq k\}$$

Одношаговая оптимальность

$$\omega_k \rightarrow \Phi(\omega_k) \rightarrow S(\omega_k) \subseteq S$$

$V_{k+1}(\omega_k, y, z)$ - эффективность проведения очередного испытания в точке (y, z) с получением результата z

Модель алгоритма

$$s = \langle \{G_k\}, \{E_k\}, \{H_k\}, \{V_k\} \rangle$$

Множество таких алгоритмов обозначим через \mathcal{S}

$$W_{k+1}(y^{k+1}) = \sup_{\varphi \in \Phi(\omega_k)} V_{k+1}(\omega_k, y^{k+1}, \varphi(y^{k+1}))$$

$$W_{k+1}(y^{k+1}) = \min_{\tilde{x}^{k+1} \in Q} W_k(\tilde{y}^{k+1}) - \text{критерий одношаговой оптимальности}$$

$$W_{k+1}(y^{k+1}) = \int_{\Phi(\omega_k)} V_{k+1}(\omega_k, y^{k+1}, z^{k+1}) dP(z^{k+1} / \omega_k, y^{k+1})$$

$P(z^{k+1} / \omega_k, y^{k+1})$ - условное (по отношению к результатам испытаний) распределение вероятностей

Асимптотическая ОПТИМАЛЬНОСТЬ

S_K - множество методов, в которых остановка осуществляется ровно K через шагов поиска

$$W(K) = \inf_{s \in S_K} W(s),$$

Если метод $s_K^* \in S_K$ - оптимальный, то очевидно, что $W(s_K^*) = W(K)$

Рассмотрим алгоритм s_∞ такой, что на каждом шаге его усечение

$$s_K \in S_K$$

Алгоритм s_∞ называется асимптотически оптимальным,

если

$$\lim_{K \rightarrow \infty} W(s_K) / W(K) = 1$$

Характеристические алгоритмы ОПТИМИЗАЦИИ

$$\varphi(x) \rightarrow \min, x \in Q = [a, b] \quad (*)$$

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

1. Задать множество

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

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 $x^i \in \Lambda_k, 1 \leq i \leq k$,

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

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

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

Характеристические алгоритмы ОПТИМИЗАЦИИ

2. Каждому интервалу (x_{i-1}, x_i) $1 \leq i \leq \tau$, 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.

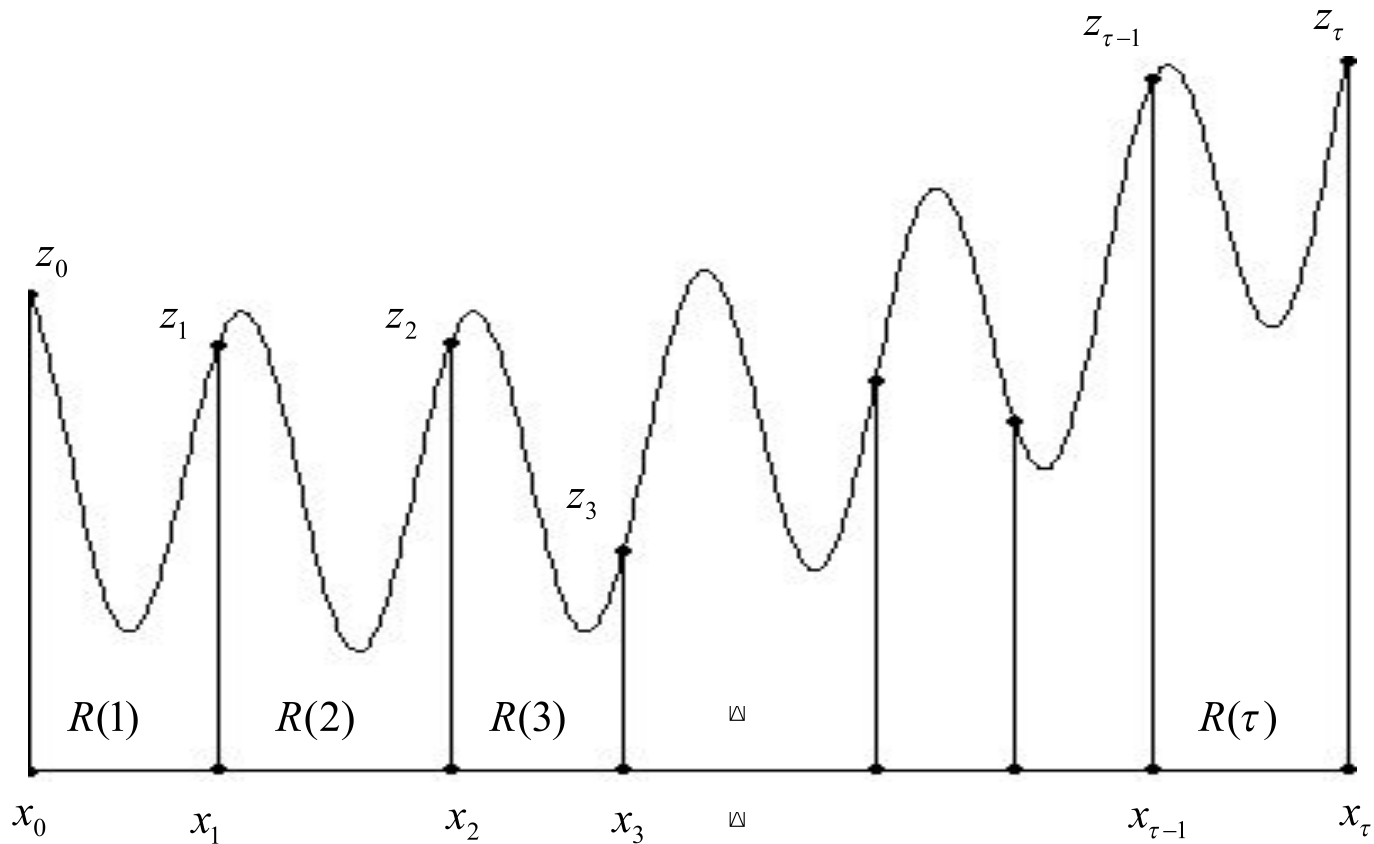
$$R(t) = \max \{R(i) : 1 \leq i \leq \tau\}$$

4. Провести очередное испытание в точке

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

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

Характеристические алгоритмы ОПТИМИЗАЦИИ



Примеры характеристических алгоритмов

Два первых испытания проводятся в точках $x^1 = a$ и $x^2 = b$,

характеристическое правило вступает в действие, начиная с $k=2$, при этом множество Λ_k ($k \geq 2$) состоит только из точек испытаний, т.е.

$\Lambda_k = \{x^1, x^2, \dots, x^k\}$ и, следовательно, $\tau = k - 1$.

Метод последовательного сканирования (перебор).

$$R(i) = x_i - x_{i-1}$$

$$x^{k+1} = 0.5(x_{i-1} + x_i).$$

Методы Пиявского и Стронгина

Метод Пиявского (метод ломаных)

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

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

Алгоритм глобального поиска – АГП (метод Стронгина)

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

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

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

$r > 1$ - параметр метода

Метод Кушнера

Характеристика подынтервала (x_{i-1}, x_i)

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

Точка очередного испытания

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

$\delta_k > 0$ - параметр метода, зависящий в общем случае от номера k испытания

$\varphi_k^* = \min_{0 \leq i \leq \tau} z_i$ - минимальное вычисленное значение после k испытаний

Алгоритм глобального поиска без вычислений на концах интервала

Согласно алгоритму одно или несколько начальных испытаний выполняются в произвольных внутренних точках отрезка $[a, b]$ и только потом вступает в действие характеристическое правило. Множество $\{x_j\}_{j=1}^k$ включает в себя координаты проведенных испытаний и дополнительно концы отрезка $a = \varphi(x_0) = \varphi(a)$ и $b = \varphi(x_{\tau}) = \varphi(b)$, однако, только для $1 \leq j \leq k$ значения $z_j = \varphi(x_j)$ и x_j не определены, поскольку точки x_0 и x_{τ} являются точками испытаний и, следовательно, значения целевой функции в них не вычисляются.

Характеристика подынтервала (x_{i-1}, x_i) вычисляется согласно выражению

$$R(i) = \begin{cases} m(x_i - x_{i-1}) + \frac{r(z_i - z_{i-1})^2}{M(x_i - x_{i-1})} - 2(z_i + z_{i-1}), & 1 < i < \tau, \\ m(x_1 - x_0) - 4z_1, & i = 1, \\ m(x_{\tau} - x_{\tau-1}) - 4z_{\tau-1}, & i = \tau, \end{cases}$$

Алгоритм глобального поиска без вычислений на концах интервала

Точка очередного испытания

$$x^{k+1} = 0.5(x_t + x_{t-1}) - \begin{cases} 0, & t = 1 \text{ or } t = \tau, \\ (z_t - z_{t-1})/(2m), & 2 \leq t \leq \tau - 1, \end{cases}$$

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

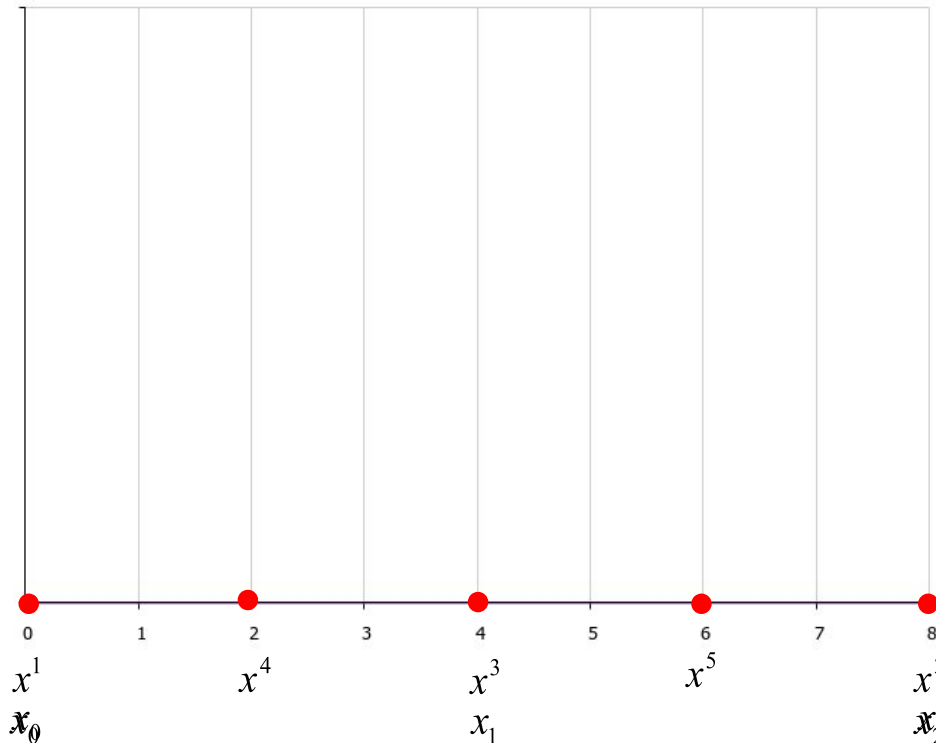
$$M = \max_{2 \leq i \leq \tau-1} \frac{|z_i - z_{i-1}|}{x_i - x_{i-1}}$$

$r > 1$ - параметр метода

Построение последовательности ИСПЫТАНИЙ

Возьмем самый простой метод – последовательного сканирования и посмотрим, как он себя ведет при оптимизации функции на отрезке $[0,8]$. Заметим, что его поведение вообще не зависит от целевой функции:

$$R(i) = x_i - x_{i-1} \quad x^{k+1} = 0.5(x_{t-1} + x_t).$$



Построение последовательности испытаний метода Стронгина

$$\varphi(x) = x^2 \rightarrow \min, x \in Q = [-4, 2]$$

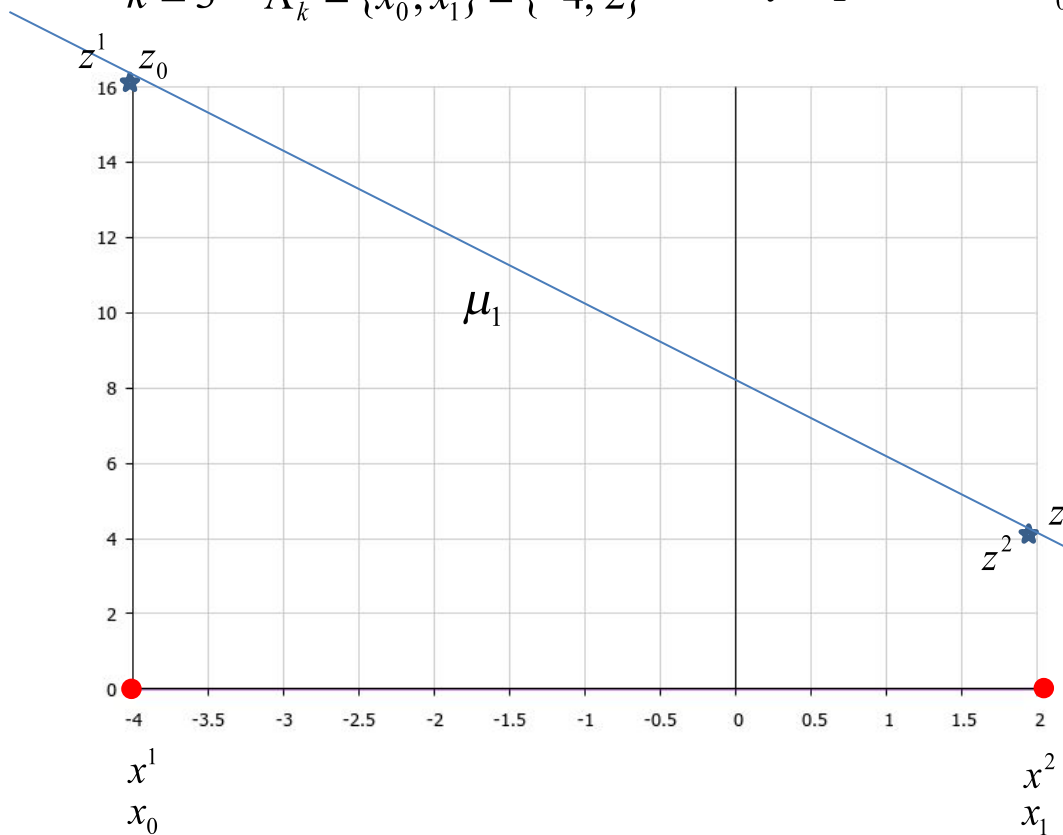
Возьмем параметр $r=2$

$$k=1 \quad x^1 = -4, \quad z^1 = \varphi(-4) = 16$$

$$k=2 \quad x^2 = 2, \quad z^2 = \varphi(2) = 4$$

$$k=3 \quad \Lambda_k = \{x_0, x_1\} = \{-4, 2\} \quad \tau = 1$$

$$z_0 = \varphi(x_0) = 16, \quad z_1 = \varphi(x_1) = 4$$



$$\mu_1 = \frac{|z_1 - z_0|}{x_1 - x_0} = \frac{12}{6} = 2$$

$$M = \max_{1 \leq i \leq \tau} \frac{|z_i - z_{i-1}|}{x_i - x_{i-1}} = \mu_1 = 2$$

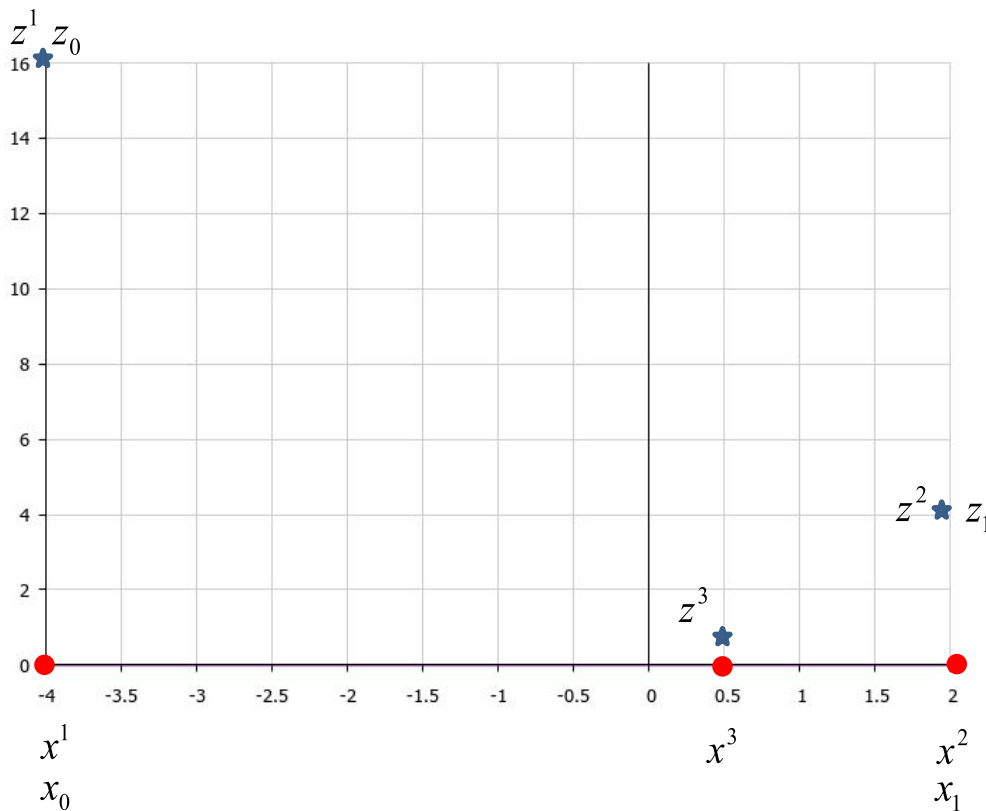
$$m = rM = 4$$

Построение последовательности испытаний метода Стронгина

$$(x_0, x_1) \rightarrow R(1) = m(x_1 - x_0) + \frac{(z_1 - z_0)^2}{m(x_1 - x_0)} - 2(z_1 + z_0) = -10$$

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

$$t = 1 \rightarrow x^{k+1} \in (x_{t-1}, x_t) = (x_0, x_1) \quad x^3 = \frac{x_1 + x_0}{2} - \frac{z_1 - z_0}{2m} = \frac{1}{2}$$



$$z^3 = \varphi(x^3) = \frac{1}{4}$$

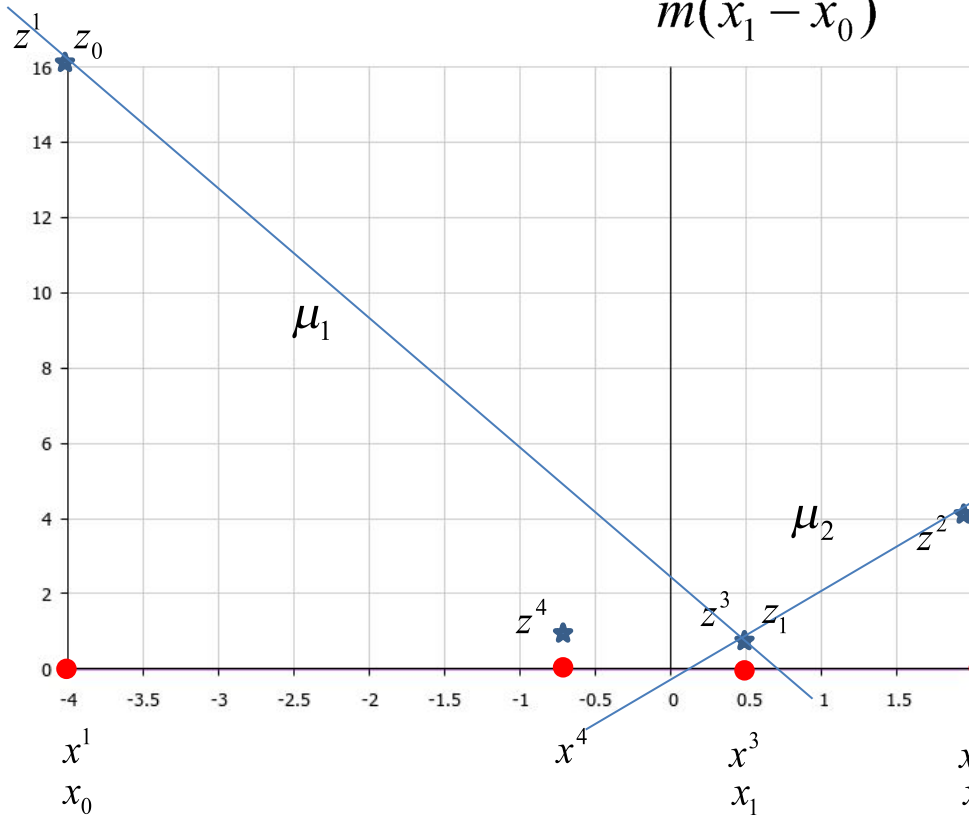
Построение последовательности испытаний метода Стронгина

$$k = 4 \quad \Lambda_k = \{x_0, x_1, x_2\} = \{-4, 0.5, 2\} \quad \tau = 2 \quad z_0 = \varphi(x_0) = 16, z_1 = \varphi(x_1) = 0.25, z_2 = \varphi(x_2) = 4$$

$$\mu_1 = \frac{|z_1 - z_0|}{x_1 - x_0} = \frac{15.75}{4.5} = 3.5 \quad \mu_2 = \frac{|z_2 - z_1|}{x_2 - x_1} = \frac{3.75}{1.5} = 2.5$$

$$M = \max_{1 \leq i \leq \tau} \frac{|z_i - z_{i-1}|}{x_i - x_{i-1}} = \mu_1 = 3.5 \quad m = rM = 7$$

$$(x_0, x_1) \rightarrow R(1) = m(x_1 - x_0) + \frac{(z_1 - z_0)^2}{m(x_1 - x_0)} - 2(z_1 + z_0) = 6.875$$



$$(x_1, x_2) \rightarrow R(2) = m(x_2 - x_1) + \frac{(z_2 - z_1)^2}{m(x_2 - x_1)} - 2(z_2 + z_1) = 3.34$$

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

$$t = 1 \rightarrow x^{k+1} \in (x_{t-1}, x_t) = (x_0, x_1)$$

$$x^4 = \frac{x_1 + x_0}{2} - \frac{z_1 - z_0}{2m} = -0.625$$

$$z^4 = \varphi(x^4) = 0.390625$$