2. ЛЕКЦИЯ. Методы нулевого порядка. Метод Нелдера-Мида.

В 1964 году Нелдер и Мид предложили модификацию, в которой симплекс может изменять свою форму (растягиваясь и сжимаясь) в зависимости от свойств поверхности целевой функции. Так как в этом случае симплекс не будет уже регулярным, метод назвали ***поиском по деформируемому многограннику.***

И так модифицируем рассмотренный на предыдущей лекции алгоритм минимизации целевой функции по регулярному симплексу, добавив к процедуре отражения при построении нового симплекса процедуры сжатия и растяжения. Геометрическая иллюстрация этих процедур для случая $n = 2$ представлена на рис. 1, 2 и 3, где введены следующие обозначения:

f_h – наибольшее значение целевой функции;

f_s – следующее по величине за наибольшим значение целевой функции;

f_l – наименьшее значение целевой функции;

$f, f^/$– текущие значения целевой функции

a) Операция отображения

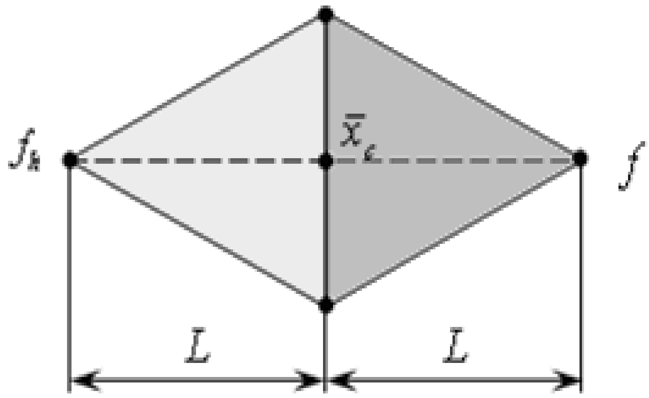


Рис.1.

b) Если $f < f_l$, то выполняется операция растяжения $L_H = \beta L$, где β – параметр растяжения.

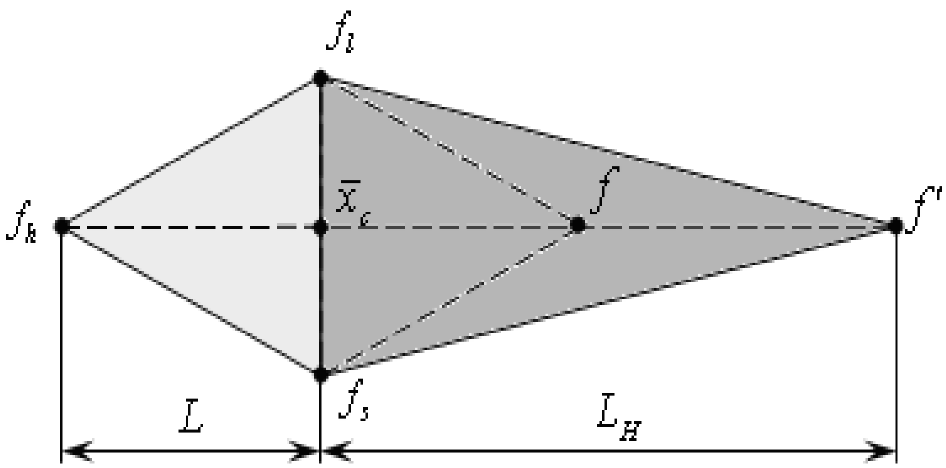


Рис.2.

c) Если $f_s < f < f_h$, то выполняется операция растяжения $L_H = \gamma L$, где γ – параметр сжатия.

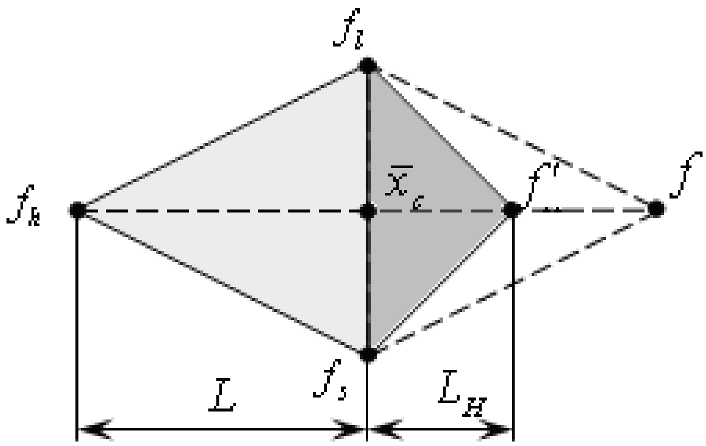


Рис.3.

При решении практически задач минимизации параметры растяжения β и сжатия γ Нелдер и Мид рекомендует брать $\beta = 2, \gamma = 0{,}5$, но Павиани – выбирать эти параметры из интервалов $2{,}8 \le \beta \le 3{,}0$ и $0{,}4 \le \gamma \le 0{,}6$

Алгоритм поиска методом Нелдера-Мида

1. Задать размерность задачи оптимизации n, координаты начальной точки многогранника $\overline{x^{(0)}} = \left\{ x_1^{(0)},\ x_2^{(0)}, \dots, x_n^{(0)} \right\}$, длину ребра многогранника m, параметр растяжения β, параметр сжатия γ (гамма), точность поиска ε.

2. Построить начальный многогранник в виде регулярного симплекса, вычисляя координаты остальных n вершин $\overline{x^{(1)}}, \overline{x^{(2)}}, \dots, \overline{x^{(n)}}$ по формулам.

$$\overline{x^{(i)}} = \begin{cases} x_j^{(0)} + \delta_1, \text{если } i = j \\ x_j^{(0)} + \delta_2, \text{если } i \ne j \end{cases} \text{ для } i, j = \overline{1, n}$$

где приращения δ_1 и δ_2 определяются по формулам

$$\delta_1 = \left(\frac{\sqrt{n+1} - 1}{n\sqrt{2}}\right) * m, \delta_2 = \left(\frac{\sqrt{n+1} + n - 1}{n\sqrt{2}}\right) * m.$$

3. Определить номер вершины k с наибольшим значением целевой функции $f_h = f(\overline{x^{(k)}})$, номер вершины k_1 с наименьшим значением целевой функции $f_l = f(\overline{x^{(k_1)}})$ и номер вершины k_2 – со следующим по величине за наибольшим значением целевой функции $f_s = f(\overline{x^{(k_2)}})$.

4. Определить центр тяжести всех вершин многогранника за исключением вершины $\overline{x^{(k)}}$:

$$\overline{x_c} = \frac{1}{n} \sum_{\substack{i=0 \\ i \ne k}}^{n} \overline{x^{(i)}}$$

5. Отразить вершину $\overline{x^{(k)}}$ относительно центра тяжести $\bar{x} = 2\overline{x_c} - \overline{x^{(k)}}$. Вычислить значение целевой функции в отраженной точке $f(\bar{x})$ и перейти к пункту 6.

6. Проверить условие. Если $f(\bar{x}) < f(\overline{x^{(k)}})$, то операция отражения закончилась успешно. Положить $\overline{x^{(k)}} = \bar{x}$, $f\left(\overline{x^{(k)}}\right) = f(\bar{x})$ и перейти к пункту 7. В противном случае перейти к пункту 9 и выполнить операцию сжатия.

7. Проверить условие. Если $f(\overline{x^{(k)}}) < f_l$, то выполнить операцию растяжения $\bar{x} = \bar{x}_c + \beta\left(\overline{x^{(k)}} - \overline{x_c}\right)$, вычислить значение целевой функции $f(\bar{x})$ и перейти к пункту 8, иначе – к пункту 9.

8. Проверить условие. Если $f(\bar{x}) < f(\overline{x^{(k)}})$, то операция растяжения закончилась успешно. Положить $\overline{x^{(k)}} = \bar{x}$, $f\left(\overline{x^{(k)}}\right) = f(\bar{x})$ и перейти к пункту 12, иначе – к пункту 9.

9. Проверить условие. Если $f_s < f(\bar{x}) < f_h$, то выполнить операцию сжатия $\bar{x} = \bar{x}_c + \gamma\left(\overline{x^{(k)}} - \overline{x_c}\right)$ вычислить значение целевой функции f((x) и перейти к пункту 10, иначе – к пункту 11.

10. Проверить условие. Если $f(\bar{x}) < f(\overline{x^{(k)}})$, то операция сжатия закончилась успешно. Положить $\overline{x^{(k)}} = \bar{x}$, $f\left(\overline{x^{(k)}}\right) = f(\bar{x})$ и перейти к пункту 12, иначе – к пункту 11.

11. Выполнить операцию редукции. Для этого определить номер вершины r с минимальным значением целевой функции $f_{min} = f\left(\overline{x^{(r)}}\right)$. Используя соотношение

$$\overline{x^{(i)}} = \overline{x^{(r)}} + 0{,}5 \cdot (\overline{x^{(i)}} - \overline{x^{(r)}}), \quad i = \overline{0,n}, \; i \neq r$$

сформировать новый многогранник с уменьшенными вдвое сторонами. Перейти к шагу 12.

12. Проверить критерий окончания процесса поиска, предложенный Нелдером и Мидом

$$\sigma = \left\{ \frac{1}{n+1} \sum_{i=0}^{n} \left[f(\overline{x^{(i)}}) - f(\overline{x_c}) \right]^2 \right\}^{\frac{1}{2}} < \varepsilon,$$

где $\overline{x_c} = \frac{1}{n+1} \sum_{i=0}^{n} \overline{x^{(i)}}$ - центр тяжести многогранника на данном шаге

13. Если условие выполнено $\sigma(\text{сигма}) < \varepsilon$, то процесс вычислений завершен. В качестве приближенного решения принять вершину многогранника с

минимальным значением целевой функции. В противном случае перейти к шагу 3 и продолжить процесс итераций.

Пример: Найти минимум целевой функции.

$$f(\overline{x}) = x_1^2 - x_1x_2 + 3x_2^2 - x_1$$

методом Нелдера-Мида с точностью ε = 0,1.

Решение. Зададим начальную точку симплекса $\overline{x^{(0)}} = (x_1^{(0)}, x_2^{(0)})^{\mathrm{T}} = (0,0)^{\mathrm{T}}$ и длину ребра симплекса m = 1, параметр растяжения $\beta = 2{,}8$, параметр сжатия $\gamma = 0{,}4$

Вычислим приращения

$\delta_1 = \left(\frac{\sqrt{n+1}-1}{n\sqrt{2}}\right) * m = \left(\frac{\sqrt{3}-1}{2\sqrt{2}}\right) *1 = 0{,}259,$

$\delta_2 = \left(\frac{\sqrt{n+1}+n-1}{n\sqrt{2}}\right) * m = \left(\frac{\sqrt{3}+2-1}{2\sqrt{2}}\right) * 1 = 0{,}966.$

Используя δ_1 и δ_2 , вычислим координаты двух остальных вершин симплекса

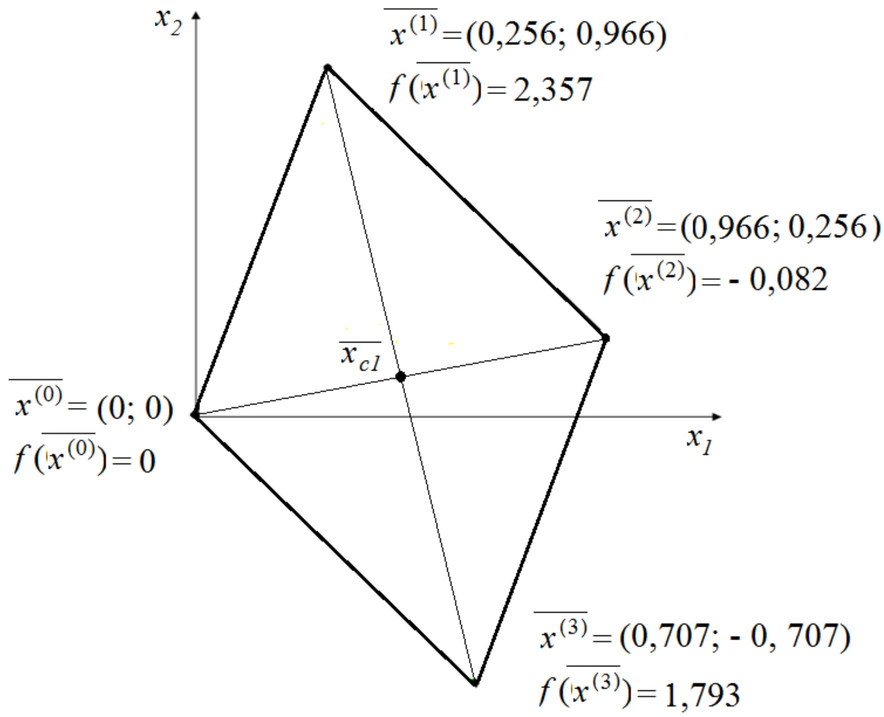
$\overline{x^{(1)}} = (x_1^{(0)} + \delta_1, x_2^{(0)} + \delta_2)^{\mathrm{T}} = (0 + 0{,}259;\ 0 + 0{,}966)^{\mathrm{T}} = (0{,}259;\ 0{,}966)^{\mathrm{T}}$

$\overline{x^{(2)}} = (x_1^{(0)} + \delta_2, x_2^{(0)} + \delta_1)^{\mathrm{T}} = (0 + 0{,}966; 0 + 0{,}259)^{\mathrm{T}} = (0{,}966;\ 0{,}259)^{\mathrm{T}}$

Итерация k = 0. Вычислим значение целевой функции $f(\overline{x})$ в вершинах $\overline{x^{(0)}}$, $\overline{x^{(1)}}$, $\overline{x^{(2)}}$ и обозначим наибольшее значение функции f_h, следующее за наибольшим значением f_s, наименьшее значение функции f_l

$f_s = f\left(\overline{x^{(0)}}\right) = 0,\ f_h = f\left(\overline{x^{(1)}}\right) = 2{,}357,\ f_l = f\left(\overline{x^{(2)}}\right) = -0{,}082.$

Составим рисунок



Составим таблицу

		Координаты вершины		Значение функции
		1	2	
Номер вершины	0	$x_1^{(0)} = 0$	$x_2^{(0)} = 0$	$f\left(\overline{x^{(0)}}\right) = 0$
	1	$x_1^{(1)} = 0{,}259$	$x_2^{(1)} = 0{,}966$	$f\left(\overline{x^{(1)}}\right) = 2{,}\ 357$
	2	$x_1^{(2)} = 0{,}966$	$x_2^{(2)} = 0{,}259$	$f\left(\overline{x^{(2)}}\right) = -0{,}082$

Наибольшее значение целевой функции соответствует вершине $\overline{x^{(1)}}$, поэтому необходимо отразить ее относительно центра тяжести остальных вершин $\overline{x^{(0)}}$ и $\overline{x^{(2)}}$. центр тяжести расположен в точке

$$\overline{x_{c1}} = \frac{1}{2}(x_1^{(0)} + x_1^{(2)}, x_2^{(0)} + x_2^{(2)})^{\mathrm{T}} = (0{,}483;\ 0{,}129)^{\mathrm{T}}.$$

Используя свойство регулярности, найдем координаты отраженной вершины

$$\overline{x^{(3)}} = 2\overline{x_{c1}} - \overline{x^{(1)}} = 2(0{,}483;\ \ 0{,}129)^{\mathrm{T}} - (0{,}259;\ \ 0{,}966)^{\mathrm{T}} = (0{,}707;\ -0{,}707)^{\mathrm{T}}.$$

В полученной вершине значение целевой функции $f\left(\overline{x^{(3)}}\right) = 1{,}793$.

		Координаты вершины		Значение функции
		1	2	
Номер вершины	0	$x_1^{(0)} = 0$	$x_2^{(0)} = 0$	$f\left(\overline{x^{(0)}}\right) = 0$
	1	$x_1^{(1)} = 0{,}259$	$x_2^{(1)} = 0{,}966$	$f\left(\overline{x^{(1)}}\right) = 2{,}\ 357$
	2	$x_1^{(2)} = 0{,}966$	$x_2^{(2)} = 0{,}259$	$f\left(\overline{x^{(2)}}\right) = -0{,}082$
	3	$x_1^{(3)} = 0{,}707$	$x_2^{(3)} = -0{,}707$	$f\left(\overline{x^{(3)}}\right) = 1{,}793$

Следовательно, наблюдается уменьшение целевой функции $f\left(\overline{x^{(3)}}\right) < f\left(\overline{x^{(1)}}\right)$

Так как выполняется условие

$$f_s = 0 < f(\overline{x^{(3)}}) = 1{,}793 < f_h = 2{,}357\,,$$

то выполним операцию сжатия симплекса.

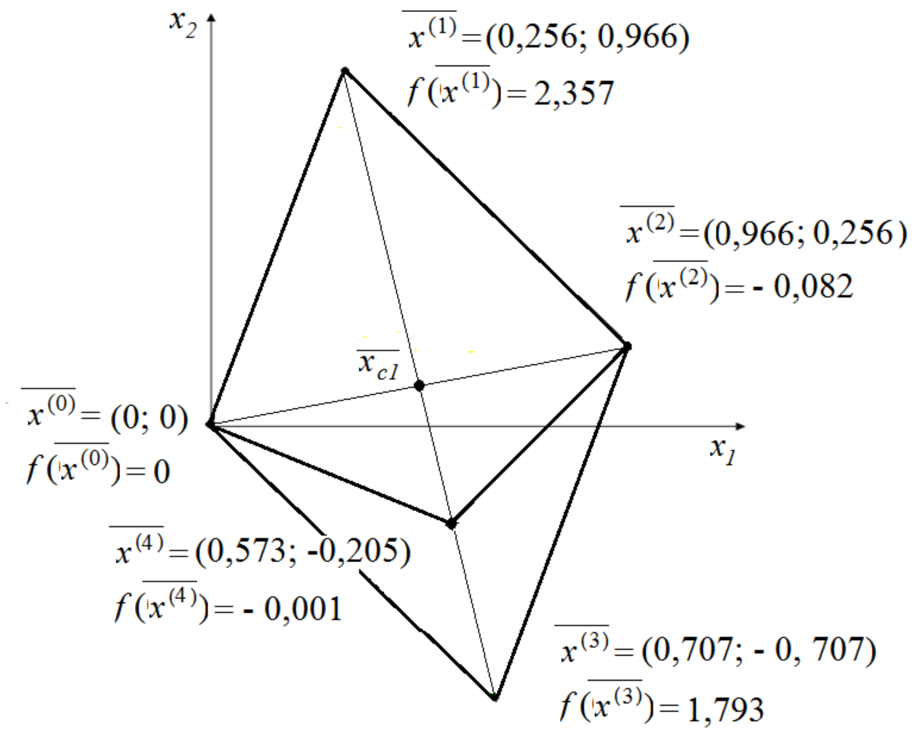
$$\overline{x^{(4)}} = \bar{x}_{c1} + 0{,}4(\overline{x^{(3)}} - \bar{x}_{c1}) =$$

$$= (0{,}483;\, 0{,}129)^T + 0{,}4[(0{,}707; -0{,}707)^T - (0{,}483;\, 0{,}129)^T] =$$

$$= (0{,}573; -0{,}205)^T.$$

В полученной вершине значение целевой функции $f\left(\overline{x^{(4)}}\right) = -0{,}001$

Так как $f\left(\overline{x^{(4)}}\right) < f\left(\overline{x^{(3)}}\right)$, то операция сжатия закончилась удачно.



Следовательно текущий симплекс образован вершинами $\overline{x^{(0)}}$, $\overline{x^{(2)}}$, $\overline{x^{(4)}}$, которым соответствуют значения целевой функции

		Координаты вершины		Значение функции
		1	2	
Номер вершины	**0**	$\boldsymbol{x_1^{(0)} = 0}$	$\boldsymbol{x_2^{(0)} = 0}$	$\boldsymbol{f(\overline{x^{(0)}}) = 0}$
	1	$x_1^{(1)} = 0{,}259$	$x_2^{(1)} = 0{,}966$	$f(\overline{x^{(1)}}) = 2{,}357$
	2	$\boldsymbol{x_1^{(2)} = 0{,}966}$	$\boldsymbol{x_2^{(2)} = 0{,}259}$	$\boldsymbol{f(\overline{x^{(2)}}) = -0{,}082}$
	3	$x_1^{(3)} = 0{,}707$	$x_2^{(3)} = -0{,}707$	$f(\overline{x^{(3)}}) = 1{,}793$
	4	$\boldsymbol{x_1^{(4)} = 0{,}573}$	$\boldsymbol{x_2^{(4)} = -0{,}205}$	$\boldsymbol{f(\overline{x^{(4)}}) = -0{,}001}$

Проверим условие окончания поиска. Определим координаты центра тяжести симплекса

$$\overline{x_c} = \frac{1}{3}(\overline{x_1^{(0)}} + \overline{x_1^{(2)}} + \overline{x_1^{(4)}};\ \overline{x_2^{(0)}} + \overline{x_2^{(2)}} + \overline{x_2^{(4)}}) =$$

$$\frac{1}{3}(0 + 0{,}966 + 0{,}573;\ 0 + 0{,}259 - 0{,}205)^T = (0{,}513; 0{,}018)^T,$$

в полученной точке

$$f(\overline{x_c}) = -0{,}258.$$

Вычислим σ *(сигма)*

$$\left\{\frac{1}{n+1}\cdot\sum_{i=0}^{2}\left[f(\overline{x^{(i)}})-f(\overline{x_c})\right]^2\right\}^{\frac{1}{2}}=$$

$$=\left\{\frac{1}{3}\left[(0+0{,}258)^2+(-0{,}082+0{,}258)^2+(-0{,}001+0{,}258)^2\right]\right\}^{\frac{1}{2}}=$$

$$=0{,}234>\varepsilon$$

Так как условие окончания поиска не выполняется, то процесс итерации должен быть продолжен.

Итерация k = 1. Текущий симплекс образован вершинами $\overline{x^{(0)}}$, $\overline{x^{(2)}}$, $\overline{x^{(4)}}$, обозначим наибольшее значение функции f_h, следующее за наибольшим значением f_s, наименьшее значение функции f_l

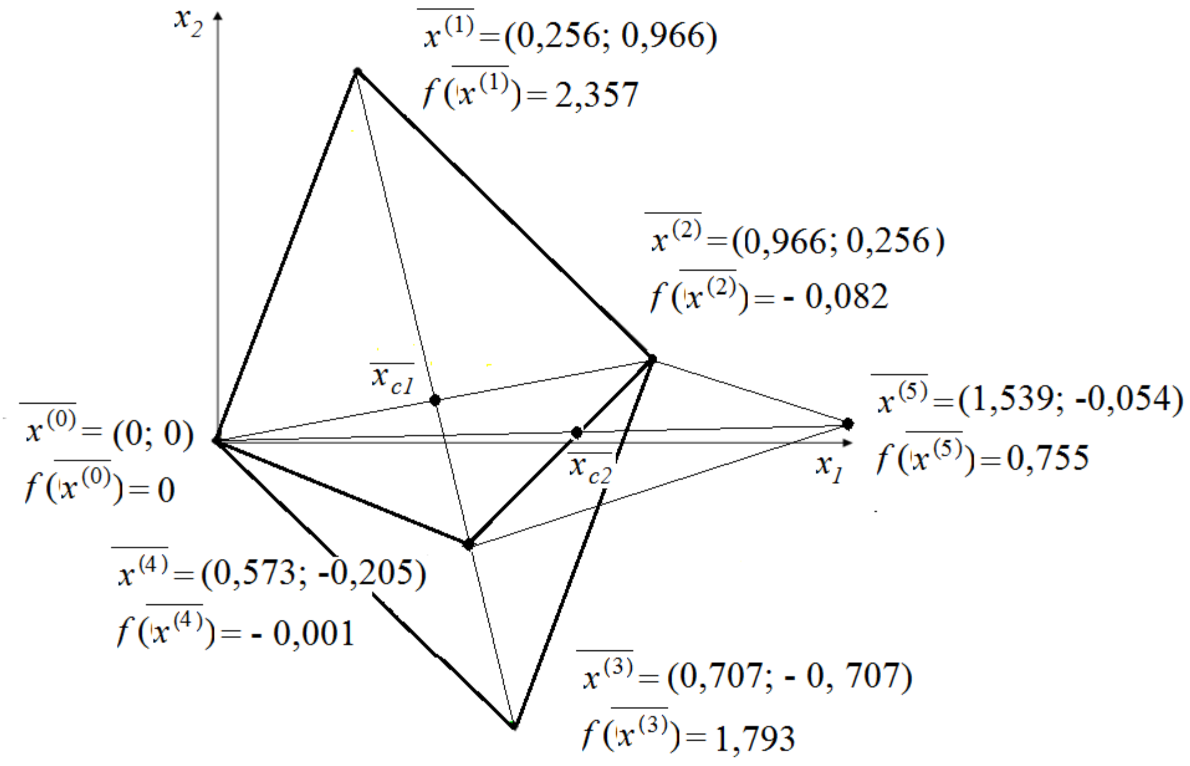
$$f_h = f\left(\overline{x^{(0)}}\right) = 0,\ f_l = f\left(\overline{x^{(2)}}\right) = -0{,}082,\ f_s = f\left(\overline{x^{(4)}}\right) = -0{,}001.$$

Наибольшее значение целевой функции соответствует вершине $\overline{x^{(0)}}$, поэтому необходимо отразить ее относительно центра тяжести остальных вершин $\overline{x^{(2)}}$ и $\overline{x^{(4)}}$. центр тяжести расположен в точке

$$\overline{x_{c2}} = \frac{1}{2}(x_1^{(2)} + x_1^{(4)}, x_2^{(2)} + x_2^{(4)})^{\mathrm{T}} = (0{,}769;\ 0{,}027)^{\mathrm{T}}$$

Используя свойство регулярности, найдем координаты отраженной вершины

$$\overline{x^{(5)}} = 2\overline{x_{c2}} - \overline{x^{(0)}} = 2(0{,}769;\ \ 0{,}027)^{\mathrm{T}} - (0;\ 0)^{\mathrm{T}} = (1{,}539;\ -0{,}054)^{\mathrm{T}}.$$



В полученной вершине $f\left(\overline{x^{(5)}}\right) = 0{,}755$

		Координаты вершины		Значение функции
		1	2	
Номер вершины	0	$x_1^{(0)} = 0$	$x_2^{(0)} = 0$	$f\left(\overline{x^{(0)}}\right) = 0$
	1	$x_1^{(1)} = 0{,}259$	$x_2^{(1)} = 0{,}966$	$f\left(\overline{x^{(1)}}\right) = 2{,}\ 357$
	2	$x_1^{(2)} = 0{,}966$	$x_2^{(2)} = 0{,}259$	$f\left(\overline{x^{(2)}}\right) = -0{,}082$
	3	$x_1^{(3)} = 0{,}707$	$x_2^{(3)} = -0{,}707$	$f\left(\overline{x^{(3)}}\right) = 1{,}793$
	4	$x_1^{(4)} = 0{,}573$	$x_2^{(4)} = -0{,}205$	$f\left(\overline{x^{(4)}}\right) = -0{,}001$
	5	$x_1^{(5)} = 1{,}539$	$x_2^{(5)} = -0{,}054$	$f\left(\overline{x^{(5)}}\right) = 0{,}755$

Так как $f\left(\overline{x^{(0)}}\right) < f\left(\overline{x^{(5)}}\right)$ то операция отражения закончилась неудачей.

Учитывая, что условие сжатия $f_s < f(\overline{x^{(5)}}) < f_h$ также не выполнено, проведем операцию редукции.

Сформируем новый многогранник с уменьшенными вдвое сторонами и вершиной $\overline{x^{(2)}}$, которой соответствует наименьшее значение целевой функции $f_l = f\left(\overline{x^{(2)}}\right) = -0{,}082$.

$$\overline{x^{(6)}} = \overline{x^{(2)}} + 0{,}5 \cdot (\overline{x^{(0)}} - \overline{x^{(2)}}) =$$

$$= (0{,}966\,;0{,}259)^T + 0{,}5 \cdot \left[(0\,;0)^T - (0{,}966\,;0{,}259)^T\right] =$$

$$= (0{,}483\,;0{,}129)^T, \qquad f(\overline{x^{(6)}}) = -0{,}262,$$

Соответственно $\overline{x^{(6)}} = \overline{x_{c1}}$

$$\overline{x^{(7)}} = \overline{x^{(2)}} + 0{,}5 \cdot (\overline{x^{(4)}} - \overline{x^{(2)}}) =$$

$$= (0{,}966\,;0{,}259)^T + 0{,}5 \cdot \left[(0{,}573\,;\text{-}\,0{,}205)^T - (0{,}966\,;0{,}259)^T\right] =$$

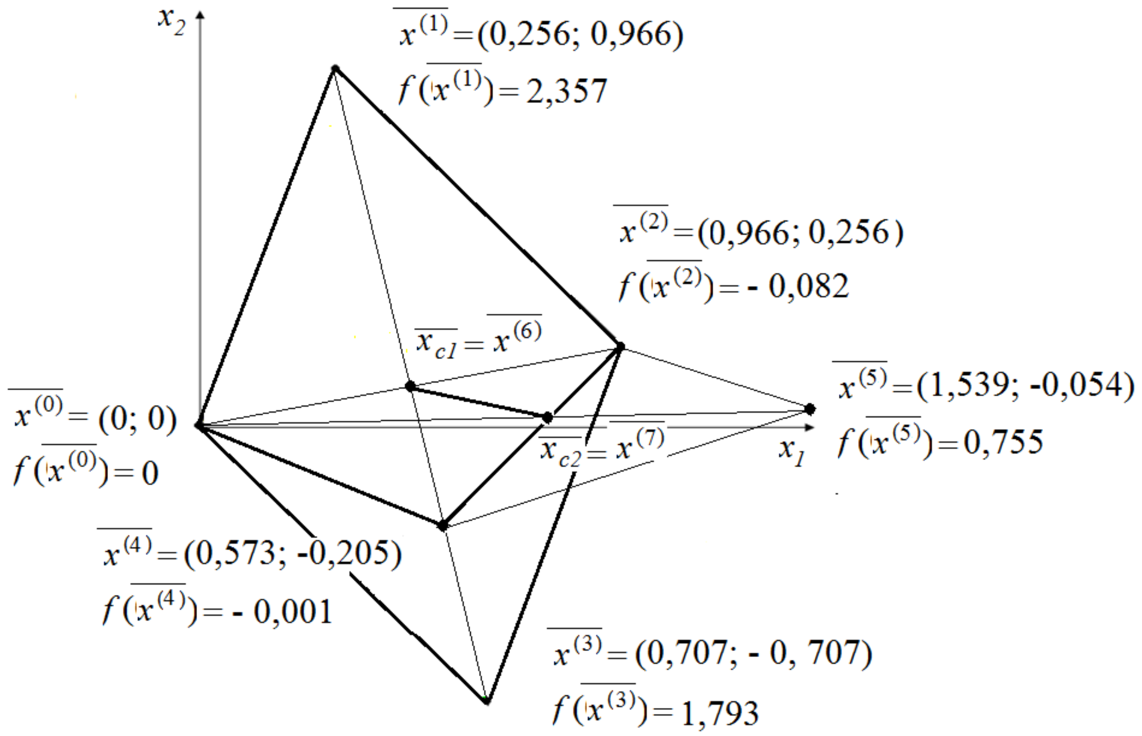
$$= (0{,}769\,;0{,}027)^T, \qquad f(\overline{x^{(7)}}) = -0{,}196.$$

Соответственно $\overline{x^{(7)}} = \overline{x_{c2}}$

		Координаты вершины		Значение функции
		1	2	
Номер вершины	0	$x_1^{(0)} = 0$	$x_2^{(0)} = 0$	$f\left(\overline{x^{(0)}}\right) = 0$
	1	$x_1^{(1)} = 0{,}259$	$x_2^{(1)} = 0{,}966$	$f\left(\overline{x^{(1)}}\right) = 2{,}\ 357$
	2	$\boldsymbol{x_1^{(2)} = 0{,}966}$	$\boldsymbol{x_2^{(2)} = 0{,}259}$	$\boldsymbol{f\left(\overline{x^{(2)}}\right) = -0{,}082}$
	3	$x_1^{(3)} = 0{,}707$	$x_2^{(3)} = -0{,}707$	$f\left(\overline{x^{(3)}}\right) = 1{,}793$

	4	$x_1^{(4)} = 0{,}573$	$x_2^{(4)} = -0{,}205$	$f\left(\overline{x^{(4)}}\right) = -0{,}001$
	5	$x_1^{(5)} = 1{,}539$	$x_2^{(5)} = -0{,}054$	$f\left(\overline{x^{(5)}}\right) = 0{,}755$
	6	$\boldsymbol{x_1^{(6)} = 0{,}483}$	$\boldsymbol{x_2^{(6)} = 0{,}129}$	$\boldsymbol{f\left(\overline{x^{(6)}}\right) = -0{,}262}$
	7	$x_1^{(7)} = 0{,}769$	$x_2^{(7)} = 0{,}027$	$f\left(\overline{x^{(7)}}\right) = -0{,}196$

Тогда.



После операции редукции текущий многогранник образован вершинами $\overline{x^{(2)}}$, $\overline{x^{(6)}}$, $\overline{x^{(7)}}$ которым соответствует значение целевой функции

$$f\left(\overline{x^{(2)}}\right) = -0{,}082,\ f\left(\overline{x^{(6)}}\right) = -0{,}262,\ f\left(\overline{x^{(7)}}\right) = -0{,}196$$

Проверим условие окончания поиска. Определим координаты центра тяжести симплекса

$$\overline{x_c} = \frac{1}{3}(\overline{x_1^{(2)}} + \overline{x_1^{(6)}} + \overline{x_1^{(7)}};\ \overline{x_2^{(2)}} + \overline{x_2^{(6)}} + \overline{x_2^{(7)}}) =$$

$$= \frac{1}{3}(0{,}966 + 0{,}483 + 0{,}769;\ 0{,}259 + 0{,}129 + 0{,}027)^T =$$

$$= (0{,}739;\ 0{,}138)^T;\quad f(\overline{x_c}) = 0{,}238.$$

Вычислим σ *(сигма)*

$$\left\{\frac{1}{n+1}\cdot\sum_{i=0}^{2}\left[f(\overline{x^{(i)}}) - f(\overline{x_c})\right]^2\right\}^{\frac{1}{2}} =$$

$$= \left\{ \frac{1}{3} \cdot \left[(-0{,}082 + 0{,}238)^2 + (-0{,}262 + 0{,}238)^2 + (-0{,}196 + 0{,}238)^2 \right] \right\}^{\frac{1}{2}} =$$

$$= 0{,}094 < \varepsilon.$$

Так как условие окончания поиска выполняется, то процесс итераций завершен.

В качестве приближенного решения $\overline{x^*}$ выбирается $\overline{x^{(6)}} = (0{,}483;\ 0{,}129)^{\mathrm{T}}$, которой соответствует наименьшее значение целевой функции $f\left(\overline{x^{(6)}}\right)$ = -0,262 текущего симплекса $\overline{x^{(2)}}$, $\overline{x^{(6)}}$, $\overline{x^{(7)}}$.