

в) метод Нелдера-Мида (деформируемого многогранника)

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 называли **поиском по деформируемому многограннику**.

Модифицируем рассмотренный в п.6 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 представлена на рис. 3.20, где введены следующие обозначения:

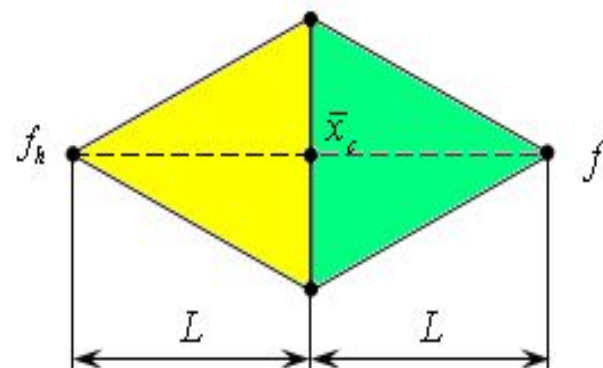
f_h - наибольшее значение целевой функции;

f_s - следующее по величине за наибольшим значение целевой функции;

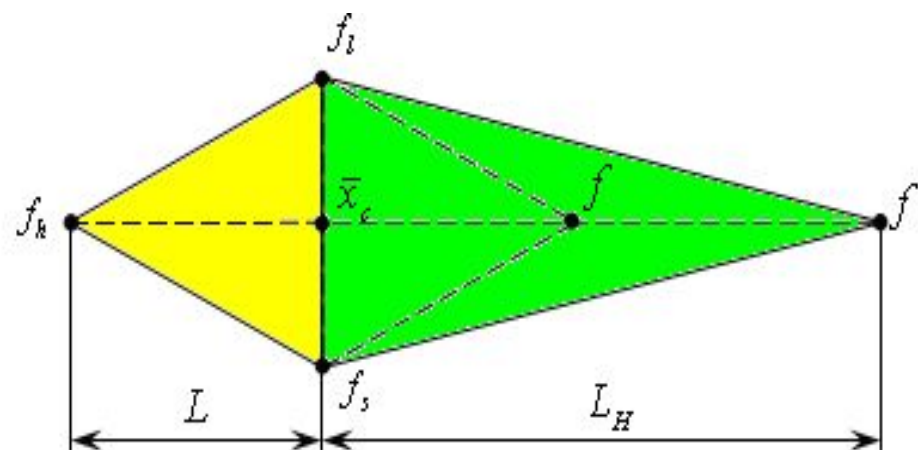
f_l - наименьшее значение целевой функции;

f, f' - текущие значения целевой функции.

а) операция отражения



б) если $f < f_l$, то выполняется операция растяжения $L_H = \beta L$,
где β - параметр растяжения



в) если $f_s < f < f_h$, то выполняется операция сжатия $L_H = \gamma L$,
 где γ - параметр сжатия

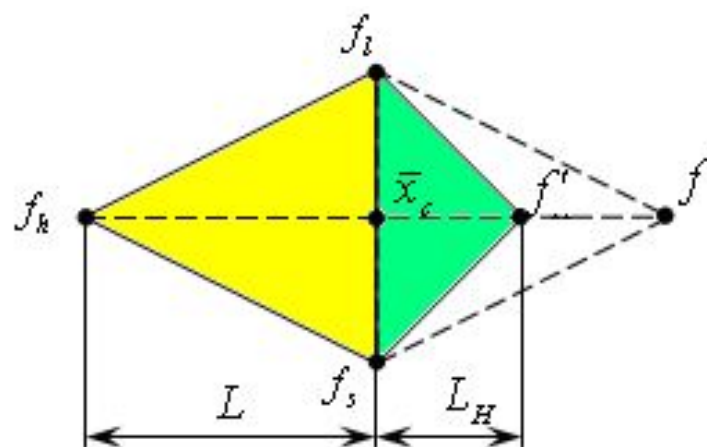


Рис. 3.20. Правила деформации симплекса

При решении практически задач минимизации параметры растяжения β и сжатия γ Нелдер и Мид рекомендует брать $\beta = 2$, $\gamma = 0,5$, Павиани – выбирать эти параметры из интервалов

$$2,8 \leq \beta \leq 3,0 ;$$

$$0,4 \leq \gamma \leq 0,6 .$$

Алгоритм поиска методом Нелдера-Мида

1. **Задать** размерность задачи оптимизации n ,
координаты начальной точки многогранника
 $\overline{x^{(0)}} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$, длину ребра многогранника
 m , параметр растяжения β , параметр сжатия
 γ , 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 (3.22).

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 $\overline{x^{(k)}}$:

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

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

6. Проверить условие. Если $f(\bar{x}) < f(\overline{x^{(k)}})$, то операция отражения закончилась успешно. Положить $\overline{x^{(k)}} = \bar{x}$, $f(\overline{x^{(k)}}) = f(\bar{x})$ 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

$$\overline{x} = \overline{x_c} + \beta \cdot (\overline{x^{(k)}} - \overline{x_c}),$$

вычислить значение целевой функции $f(\overline{x})$ и
перейти к пункту 8, иначе – к пункту 9.

8. Проверить условие. Если $f(\overline{x}) < f(\overline{x^{(k)}})$, то

операция растяжения закончилась успешно.

Положить $\overline{x^{(k)}} = \overline{x}$, $f(\overline{x^{(k)}}) = f(\overline{x})$ и перейти к
пункту 12, иначе – к пункту 9.

9. Проверить условие. Если $f_s < f(\bar{x}) < f_h$ то
выполнить операцию сжатия

$$\bar{x} = \bar{x}_c + \gamma(\bar{x}^{(k)} - \bar{x}_c)$$

вычислить значение целевой функции $f(\bar{x})$ и
перейти к пункту 10, иначе к пункту 11.

10. Проверить условие. Если $f(\bar{x}) < f(\bar{x}^{(k)})$, то
операция сжатия закончилась успешно.

Положить $\bar{x}^{(k)} = \bar{x}$, $f(\bar{x}^{(k)}) = f(\bar{x})$ и перейти к
пункту 12, иначе к пункту 11.

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

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

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 $\sigma < \varepsilon$, 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 3.9. Найти минимум целевой функции

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

методом Нелдера-Мида с точностью $\varepsilon = 0,1$.

Решение. Зададим начальную точку симплекса

$\overline{x^{(0)}} = (x_1^{(0)}, x_2^{(0)})^T = (0,0)^T$, длину ребра симплекса $m = 1$,
параметр растяжения $\beta = 2,8$, параметр сжатия $\gamma = 0,4$.
Вычислим приращения

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

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

Используя δ_1 и δ_2 вычислим координаты двух остальных вершин симплекса

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

$$\begin{aligned}\overline{x^{(2)}} &= (x_1^{(0)} + \delta_2; x_2^{(0)} + \delta_1)^T = \\ &= (0 + 0,966; 0 + 0,259)^T = (0,966; 0,259)^T .\end{aligned}$$

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

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

Наибольшее значение целевой функции соответствует вершине $\bar{x}^{(1)}$ отразим ее относительно центра тяжести вершин $\bar{x}^{(0)}$ и $\bar{x}^{(2)}$.

$$\bar{x}_{c1} = \frac{1}{2} (x_1^{(0)} + x_1^{(2)}; x_2^{(0)} + x_2^{(2)})^T = (0,483; 0,129)^T.$$

Используя свойство **регулярности**, найдем координаты отраженной вершины

$$\begin{aligned}\overline{x^{(3)}} &= 2\overline{x_{c1}} - \overline{x^{(1)}} = \\ &= 2(0,483; 0,129)^T - (0,259; 0,966)^T = \\ &= (0,707; -0,707)^T. \quad f(\overline{x^{(3)}}) = 1,793.\end{aligned}$$

Так как $f(\overline{x^{(3)}}) < f(\overline{x^{(1)}})$, то **отражение удачно.**

Условие растяжения $f(\overline{x^{(3)}}) < f_l$ **не выполняется.**

Так как выполняется условие

$f_s = 0 < f(\overline{x^{(3)}}) = 1,793 < f_h = 2,357$,
то выполним операцию **сжатия симплекса**

$$\begin{aligned}
 \overline{x^{(4)}} &= \overline{x_{c1}} + 0,4(\overline{x^{(3)}} - \overline{x_{c1}}) = \\
 &= (0,483; 0,129)^T + 0,4[(0,707; -0,707)^T - (0,483; 0,129)^T] = \\
 &= (0,573; -0,205)^T.
 \end{aligned}$$

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

Следовательно **текущий симплекс образован вершинами**

$$\begin{aligned}
 \overline{x^{(0)}}, \quad f(\overline{x^{(0)}}) &= 0, \\
 \overline{x^{(2)}}, \quad f(\overline{x^{(2)}}) &= -0,082, \\
 \overline{x^{(4)}}, \quad f(\overline{x^{(4)}}) &= -0,001.
 \end{aligned}$$

Проверим условие окончания поиска. Определим координаты центра тяжести симплекса

$$\begin{aligned}\overline{x}_c &= \frac{1}{3}(\overline{x_1^{(0)}} + \overline{x_1^{(2)}} + \overline{x_1^{(4)}}; \overline{x_2^{(0)}} + \overline{x_2^{(2)}} + \overline{x_2^{(4)}}) = \\ \frac{1}{3}(0 + 0,966 + 0,573; 0 + 0,259 - 0,205)^T &= (0,513; 0,018)^T, \\ f(\overline{x}_c) &= -0,258.\end{aligned}$$

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

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

$$= 0,234 > \varepsilon$$

Процесс итераций продолжается.

Итерация 1. Текущий симплекс образован вершинами $\overline{x^{(0)}}$, $\overline{x^{(2)}}$, $\overline{x^{(4)}}$, которым соответствует значение целевой Функции

$$f_h = f(\overline{x^{(0)}}) = 0,$$

$$f_l = f(\overline{x^{(2)}}) = -0,082,$$

$$f_s = f(\overline{x^{(4)}}) = -0,001.$$

Отразим вершину $\overline{x^{(0)}}$ относительно центра тяжести вершин $\overline{x^{(2)}}$ и $\overline{x^{(4)}}$

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

Используя свойство регулярности, найдем координаты отраженной вершины

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

$$f(\overline{x}^{(5)}) = 0,755.$$

Так как $f(\overline{x}^{(0)}) < f(\overline{x}^{(5)})$, 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 $f_l = f(\overline{x^{(2)}}) = -0,082$

$$\begin{aligned}\overline{x^{(6)}} &= \overline{x^{(2)}} + 0,5 \cdot (\overline{x^{(0)}} - \overline{x^{(2)}}) = \\ &= (0,966 ; 0,259)^T + 0,5 \cdot [(0 ; 0)^T - (0,966 ; 0,259)^T] = \\ &= (0,483 ; 0,129)^T, \quad f(\overline{x^{(6)}}) = -0,262, \\ \overline{x^{(7)}} &= \overline{x^{(2)}} + 0,5 \cdot (\overline{x^{(4)}} - \overline{x^{(2)}}) = \\ &= (0,966 ; 0,259)^T + 0,5 \cdot [(0,573 ; -0,205)^T - (0,966 ; 0,259)^T] = \\ &= (0,769 ; 0,027)^T, \quad f(\overline{x^{(7)}}) = -0,196.\end{aligned}$$

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

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

Проверим условие окончания поиска. Определим координаты центра тяжести симплекса

$$\begin{aligned} \overline{x_c} &= \frac{1}{3}(\overline{x_1^{(2)}} + \overline{x_1^{(6)}} + \overline{x_1^{(7)}}; \overline{x_2^{(2)}} + \overline{x_2^{(6)}} + \overline{x_2^{(7)}}) = \\ &= \frac{1}{3}(0,966 + 0,483 + 0,769; 0,259 + 0,129 + 0,027)^T = \\ &= (0,739; 0,138)^T; \quad f(\overline{x_c}) = 0,238. \end{aligned}$$

Вычислим

$$\begin{aligned} & \left\{ \frac{1}{n+1} \cdot \sum_{i=0}^2 \left[f(\overline{x^{(i)}}) - f(\overline{x_c}) \right]^2 \right\}^{\frac{1}{2}} = \\ & = \left\{ \frac{1}{3} \cdot [(-0,082 + 0,238)^2 + (-0,262 + 0,238)^2 + (-0,196 + 0,238)^2] \right\}^{\frac{1}{2}} = \\ & = 0,094 < \varepsilon. \end{aligned}$$

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

В качестве приближенного значения выбирается вершина с наименьшим значением целевой функции текущего симплекса, образованного вершинами $\overline{x^{(2)}}$, $\overline{x^{(6)}}$, $\overline{x^{(7)}}$:

$$\overline{x^*} = \overline{x^{(6)}} = (0,438; 0,129)^T,$$

$$f(\overline{x^*}) = -0,262.$$

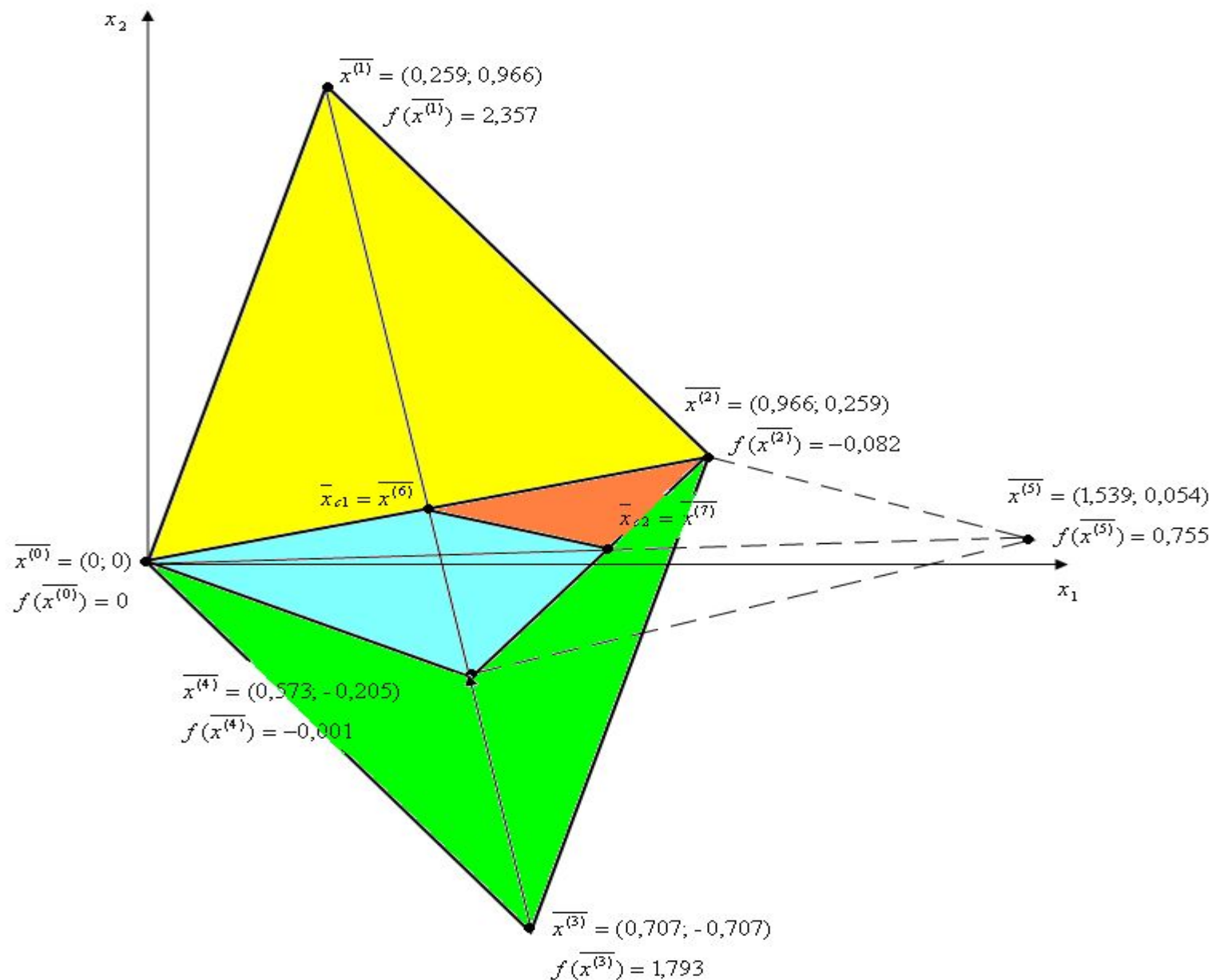


Рис. 3.21. Итерации поиска методом Нелдера - Мида в примере 3.9