Метод деформируемого многогранника (Нелдера-Мида)

Кочеганова Л.М.
М21-ИВТ-1

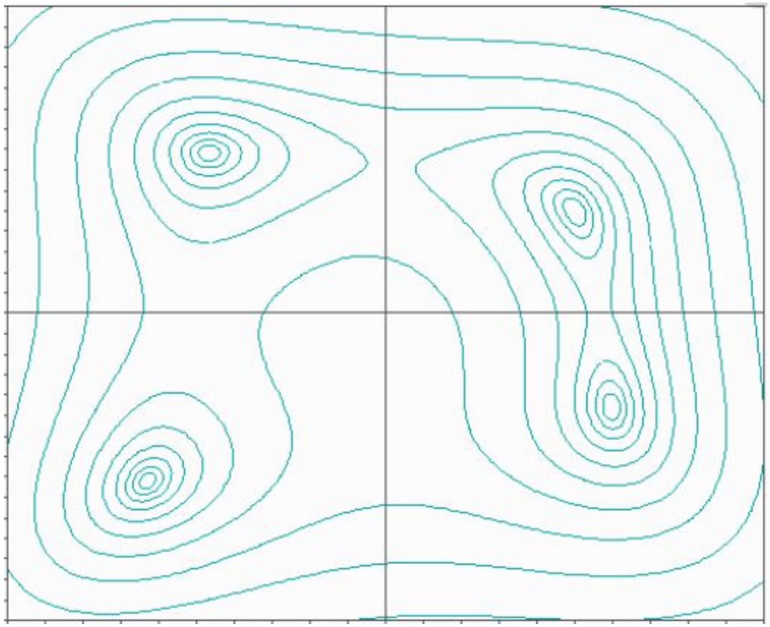
Задача

$$\min_{\mathbf{X}\in R^n} \Phi(\mathbf{X}) = \Phi(\mathbf{X}^*) = \Phi^*$$

Найти минимум критерия оптимальности $\Phi(\mathbf{X})$,определенного в n-мерном евклидовом пространстве R^n

Метод использует следующие операции над симплексами:

1. отражение;
2. редукция;
3. сжатие;
4. растяжение.



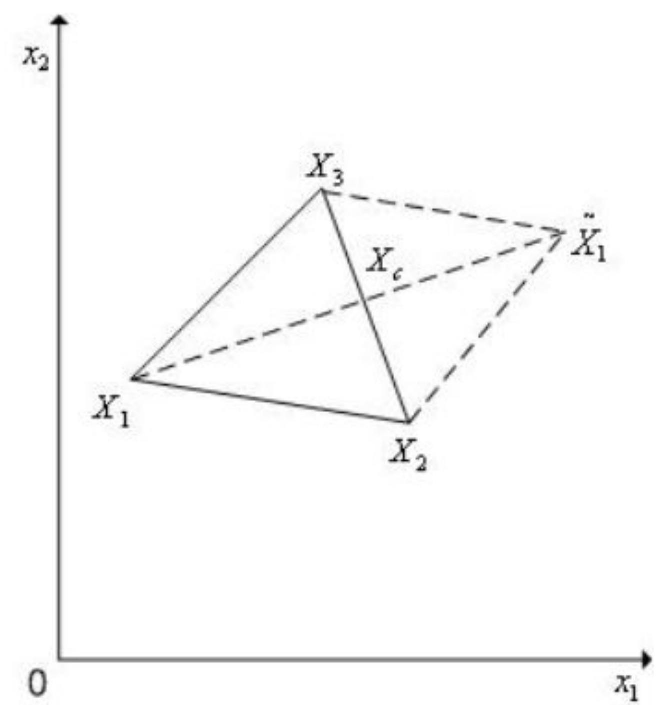
Отражение вершин симплекса

Производится относительно центра тяжести X_c остальных вершин.
Отражение k-й вершины с координатами:

$$\mathbf{X}_i^r,\ i\in[1,n+1]$$

Новый симплекс с координатами вершин:

$$\mathbf{X}_i^{r+1}=\mathbf{X}_i^r,\ i\in[1,n+1],\ i\neq k,\ \mathbf{X}_k^{r+1}=2\mathbf{X}_c^r-\mathbf{X}_k^r$$



Вектор координат центра тяжести остальных вершин симплекса:

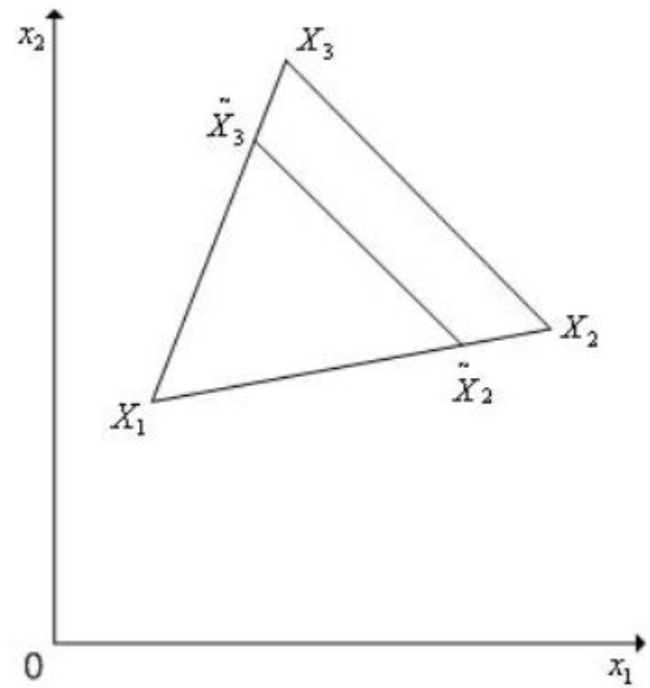
$$\mathbf{X}_c^r=\frac{1}{n}\sum_{i=1,i\neq k}^{n+1}\mathbf{X}_i^r$$

Редукция симплекса

Уменьшение длин всех ребер симплекса в одно и то же количество раз.
Редукция вершин симплекса $\mathbf{X}_i^r,\ i\in[1,n+1]$ к верши $\mathbf{X}_k$

Новый симплекс с координатами вершин:

$$\mathbf{X}_i^{r+1}=\mathbf{X}_k^r+\gamma(\mathbf{X}_i^r-\mathbf{X}_i^k),\ i\in[1,n+1],\ i\neq k,\ \mathbf{X}_k^{r+1}=\mathbf{X}_k^r.$$



Коэффициенты редукции равны:

$\gamma\in(0,\ 1),\ \ \gamma\approx0.5$

Сжатие симплекса

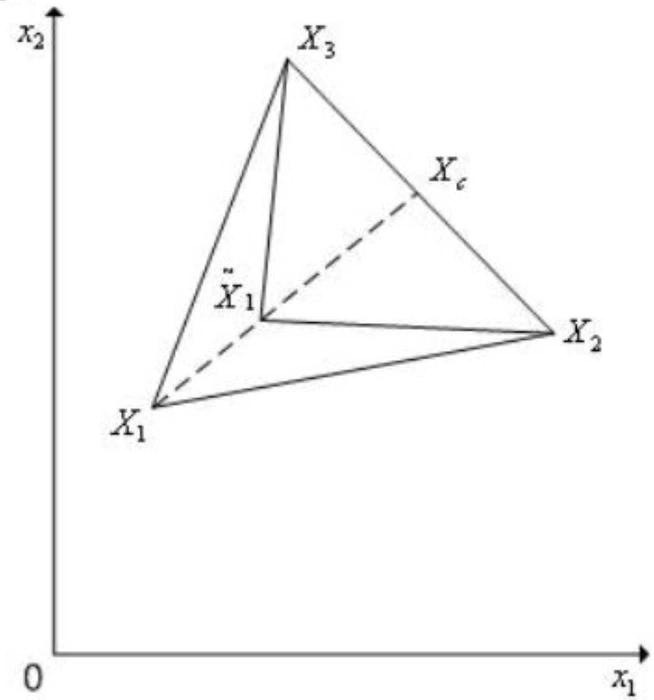
Сжатие симплекса $\mathbf{X}_i^r$, $i \in [1, n+1]$ в направлен $(\mathbf{X}_k^r - \mathbf{X}_c^r)$

Новый симплекс с координатами вершин:

$$\mathbf{X}_i^{r+1} = \mathbf{X}_i^r,\ i \in [1, n+1],\ i \neq k,\ \mathbf{X}_k^{r+1} = \mathbf{X}_c^r + \beta(\mathbf{X}_k^r - \mathbf{X}_c^r)$$

Коэффициенты сжатия:

$\beta \in (0, 1)$, $\beta \approx 0.4\text{-}0.6$



Растяжение симплекса

Сжатие симплекса $\mathbf{X}_i^r,\ i\in[1,n+1]$ в направлен $(\mathbf{X}_k^r\text{-}\mathbf{X}_c^r)$

Новый симплекс с координатами вершин:

$\mathbf{X}_i^{r+1}=\mathbf{X}_i^r,\ i\in[1,n+1],\ i\neq k,\ \mathbf{X}_k^{r+1}=\mathbf{X}_c^r+\alpha(\mathbf{X}_k^r\text{-}\mathbf{X}_c^r)$

Коэффициенты растяжения:

$\alpha\approx 2.8\text{-}3.0$

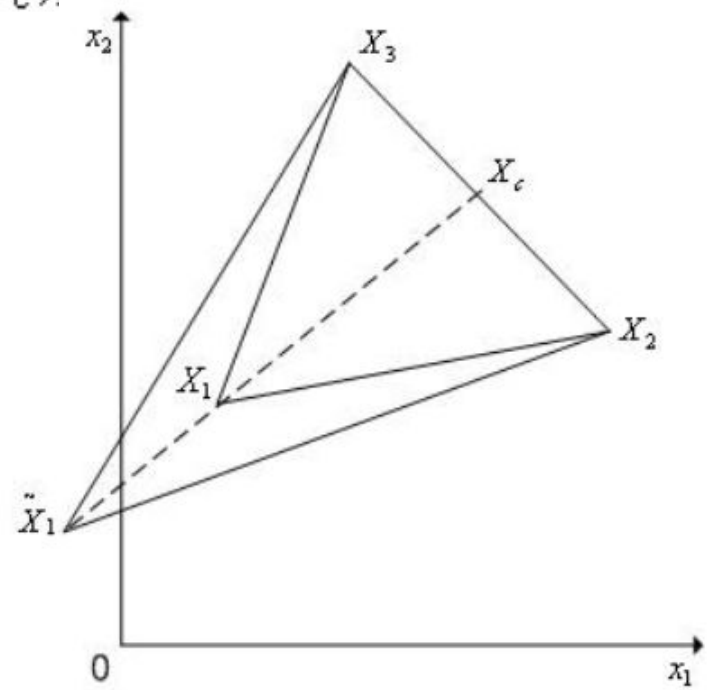


Схема метода Нелдера-Мида

$\mathbf{X}_i^r,\ i\in[1,n+1]$ обозначим за S^r; зададим начальную точку $\mathbf{X}^0$

Находим координаты $\mathbf{X}_i^0,\ i\in[1,n+1]$ и вычисляем значение функции во всех вершинах симплекса $\Phi(\mathbf{X}_i^r)$

Среди вершин симплекса найдем те, что принимают наименьшее, наибольшее и следующее за максимальным значение $\mathbf{X}_{l=k_l}^r, \mathbf{X}_{h=k_h}^r, \mathbf{X}_{s=k_s}^r$
Вычислим значение функции в них:

$$\Phi(\mathbf{X}_l^r) = \min_{i\in[1,n+1]} \Phi(\mathbf{X}_i^r),$$

$$\Phi(\mathbf{X}_h^r) = \max_{i\in[1,n+1]} \Phi(\mathbf{X}_i^r),$$

$$\Phi(\mathbf{X}_s^r) = \max_{i\in[1,n+1],\, i\neq h} \Phi(\mathbf{X}_i^r)$$

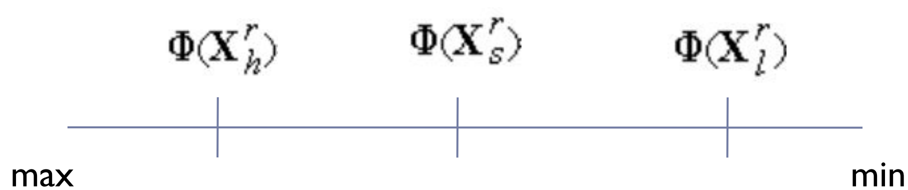


Схема метода Нелдера-Мида

Выполняем отражение вершины симплекса $\mathbf{X}_h^r$, вычисляем $\Phi(\mathbf{X}_{k_h}^{r+1})$,
получили новый симплекс S^{r+1}

Если $\Phi(\mathbf{X}_{k_h}^{r+1}) \le \Phi(\mathbf{X}_h^r)$ и $\Phi(\mathbf{X}_{k_h}^{r+1}) \le \Phi(\mathbf{X}_l^r)$.

Если $\Phi(\mathbf{X}_{k_h}^{r+1}) \le \Phi(\mathbf{X}_h^r)$ но $\Phi(\mathbf{X}_{k_h}^{r+1}) \ge \Phi(\mathbf{X}_l^r)$.

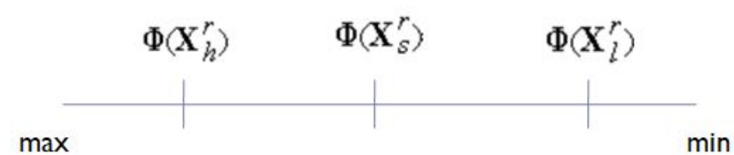
Если $\Phi(\mathbf{X}_{k_h}^{r+1}) \ge \Phi(\mathbf{X}_h^r)$

Растяжение симплекса в направлении $(\mathbf{X}_{k_h}^{r+1} - \mathbf{X}_c^{r+1})$
Получаем новую вершину $\mathbf{X}_{k_h}^{r+2}$ Если $\Phi(\mathbf{X}_{k_h}^{r+2}) \le \Phi(\mathbf{X}_l^r)$ возвращаемся к вычислению 3х вершин симплекса и отражению.
$r = r + 2$

Иначе - данные не используем.
$r = r + 1$

Возвращаемся к вычислению 3х вершин симплекса и отражению.

Сжатие симплекса в направлении $(\mathbf{X}_{k_h}^{r+1} - \mathbf{X}_c^{r+1})$
Получаем новую вершину $\mathbf{X}_{k_h}^{r+2}$. Если $\Phi(\mathbf{X}_{k_h}^{r+2}) \le \Phi(\mathbf{X}_{k_h}^r)$, то возвращаемся к вычислению 3х вершин симплекса и отражению. $r = r + 2$
Иначе – выполняем редукцию к вершине $\mathbf{X}_{l=k_l}^r$



Условие окончания итераций

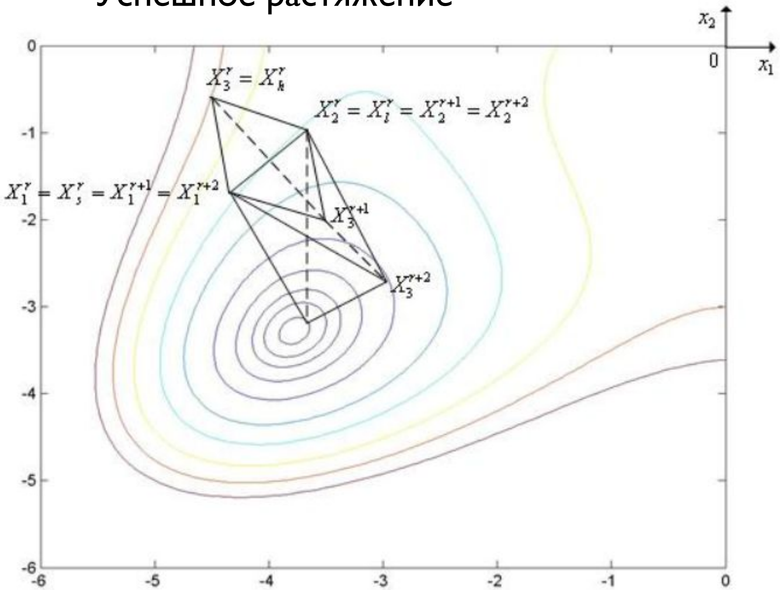
$$\max_{i\in[1,n+1],i\neq k}\left|\Phi(\mathbf{X}_i^{r+1})\text{-}\Phi(\mathbf{X}_k^{r+1})\right|\leq\varepsilon_\Phi;$$

ε_Φ - требуемая точность решения

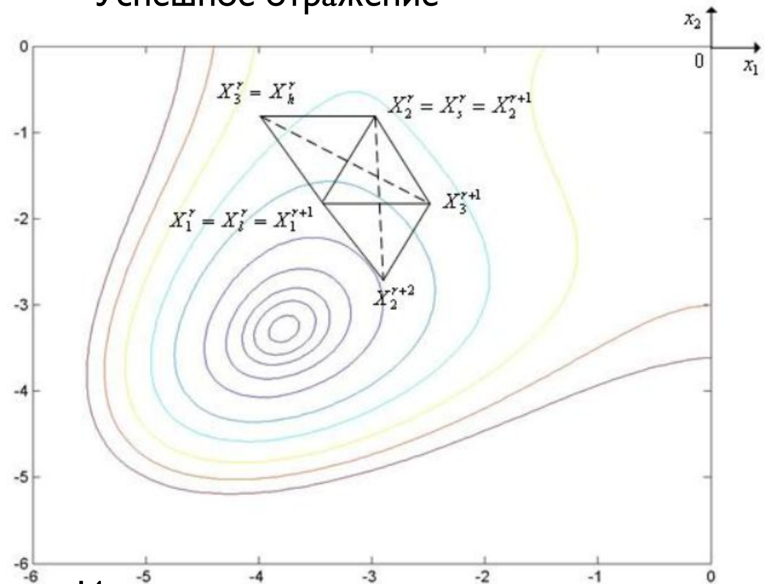
Можно завершать итерации, когда длина максимального из ребер текущего симплекса станет меньше или равна требуемой точности решения ε_X

Также можно завершить алгоритм, если выполнено необходимое количество итераций.

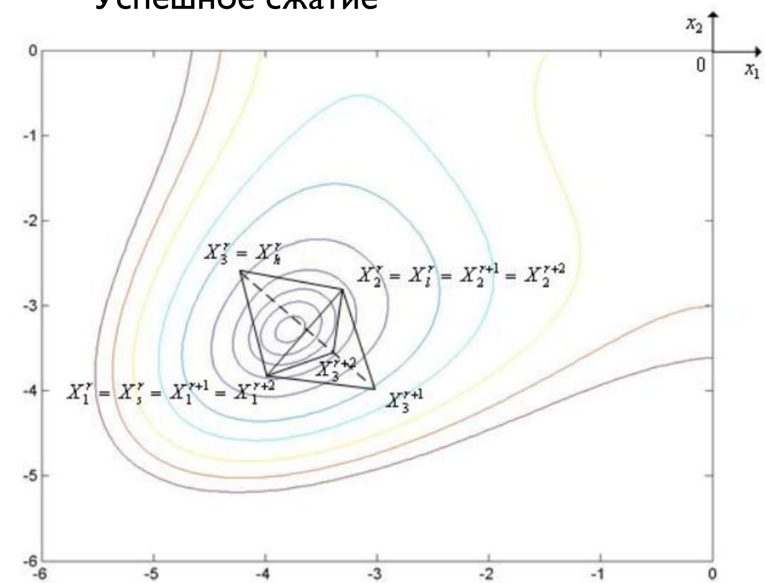
Успешное растяжение



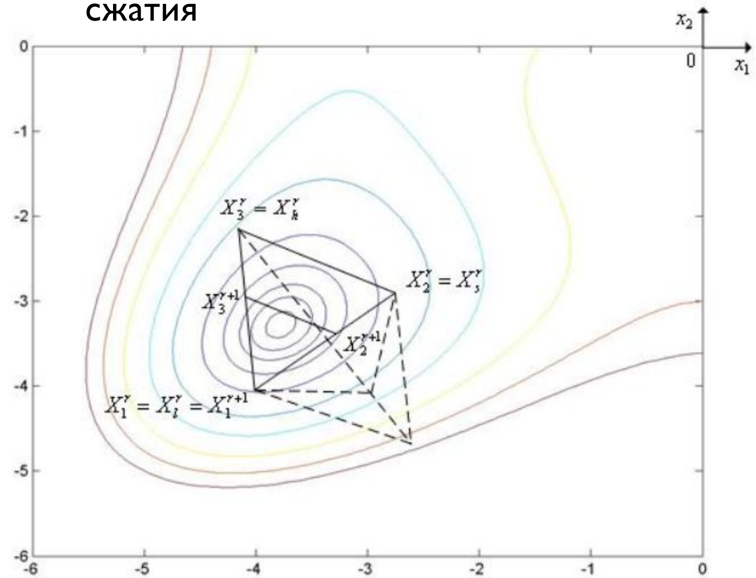
Успешное отражение



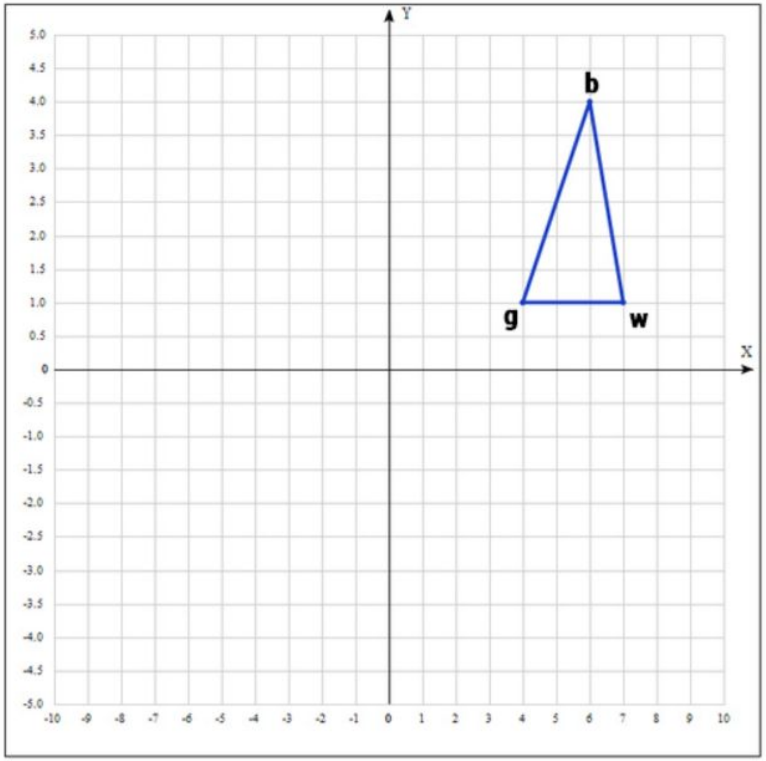
Успешное сжатие



Использование редукции после неудачного сжатия



1)



2)

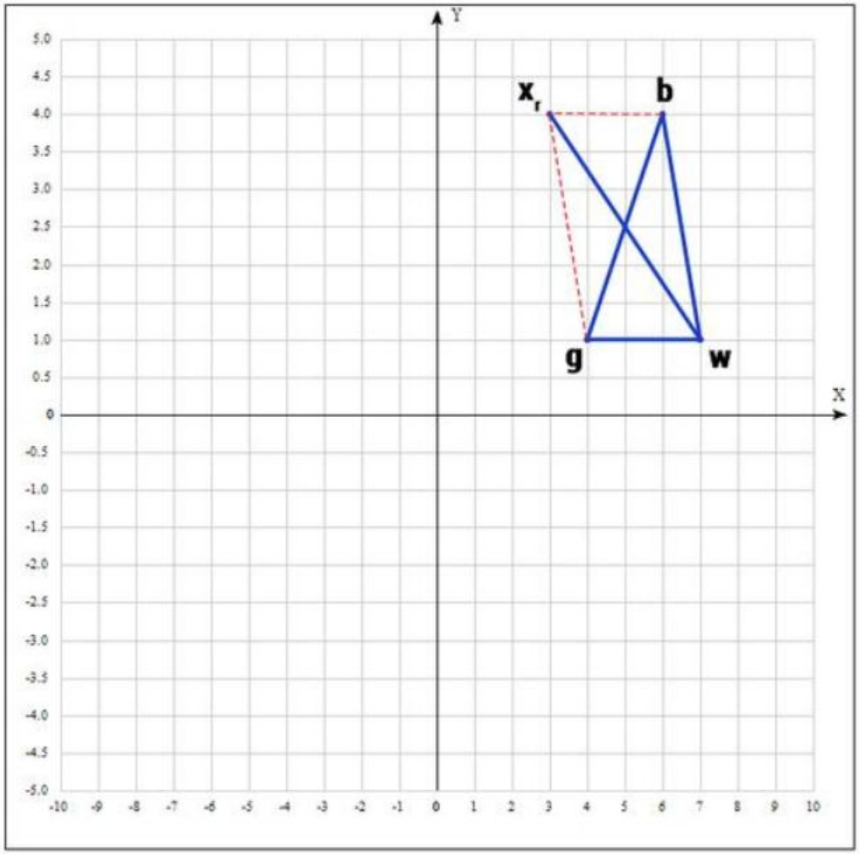
$$mid = \left(\frac{x_1 + x_2}{2}; \frac{y_1 + y_2}{2}\right)$$

$$mid = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$f(V_2) \leq f(V_1) \leq f(V_3).$$

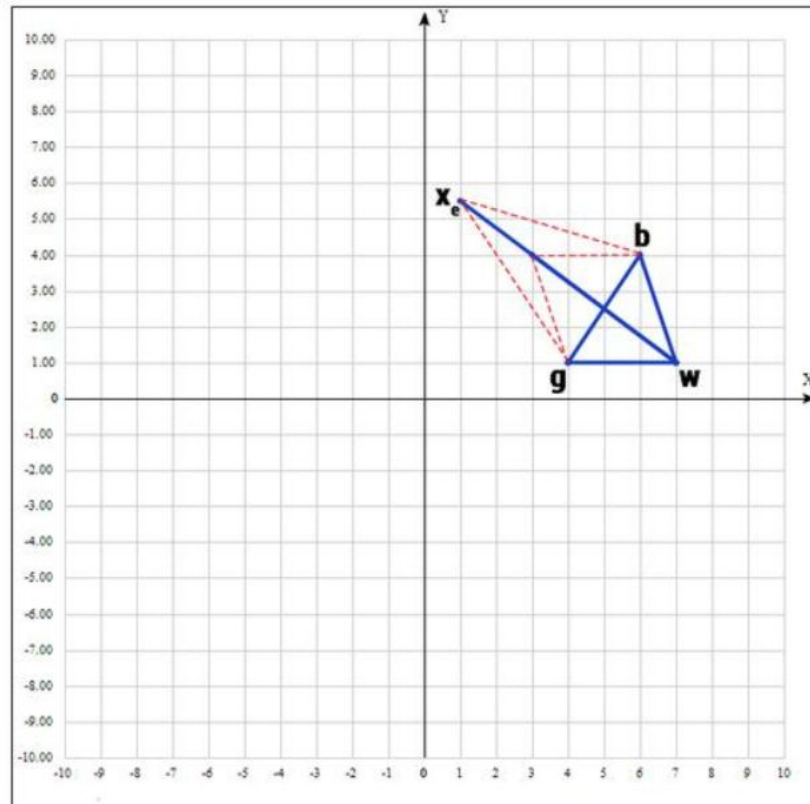
b = V_2, g = V_1, w = V_3, где best, good, worst — соответственно.

3) $x_r = mid + (mid - w)$



4) $x_e = mid + (x_r - mid)$

$f(x_e) < f(b)$



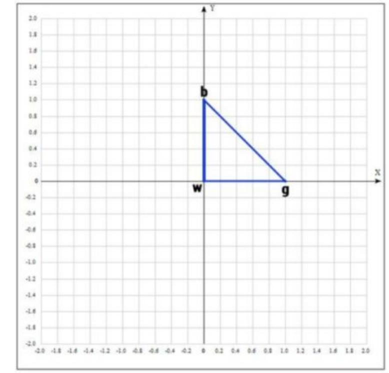
Пример

Найти экстремум следующей функции: $F(x,y)=x2+xy+y2-6x-9y$

Возьмем точки: $V1(0, 0)$; $F(V1) = 0 = b$;

$V2(1, 0)$; $F(V2) = -5 = g$;

$V3(0, 1)$; $F(V3) = -8 = w$.

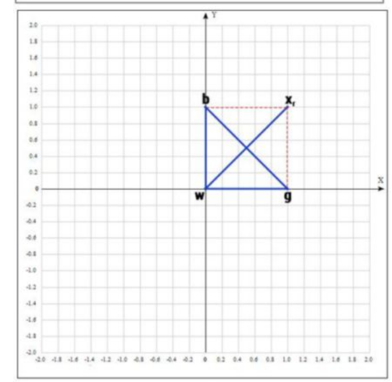


Находим центр тяжести X_c (середина отрезка bg) = $(b+g)/2 = (½, ½)$

Отражение вершины: $X1 = 2X_c - w = (1,1)$

Вычисляем значение в точке X1

$F(X1) = -12$; $F(X1) < F(b)$ ⟹ операция растяжения



Растяжение вершины: $X2 = X_c + \alpha(X1 - X_c)$

Возьмем $\alpha = 2$ ⟹ $X2 = (1.5, 1.5)$

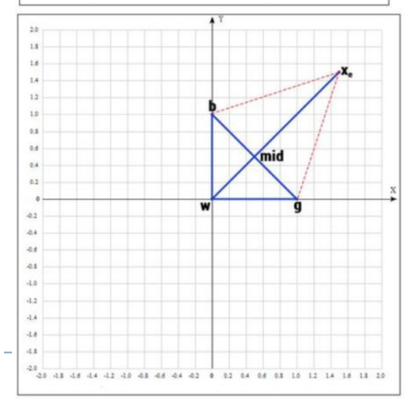
Вычисляем значение в точке X2

$F(X2) = -15.75$ ⟹ получаем новые вершины

$V1(1.5, 1.5)$; $F(V1) = -15.75$

$V2(1, 0)$; $F(V2) = -5$

$V3(0, 1)$; $F(V3) = -8$



Начинаем алгоритм сначала!

Пример

Результат для 10 итераций

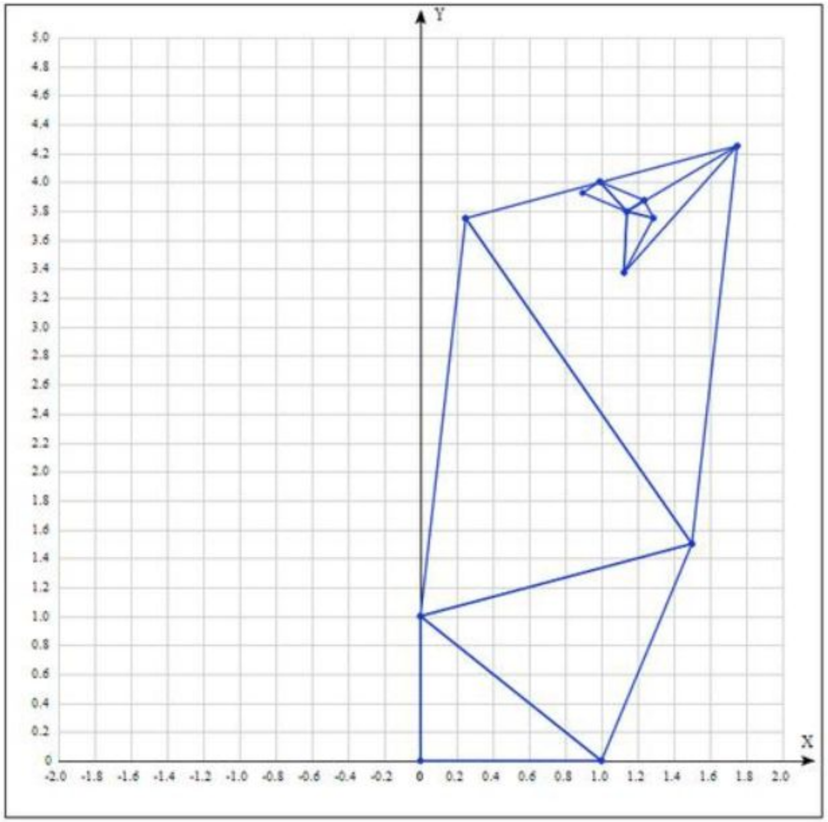
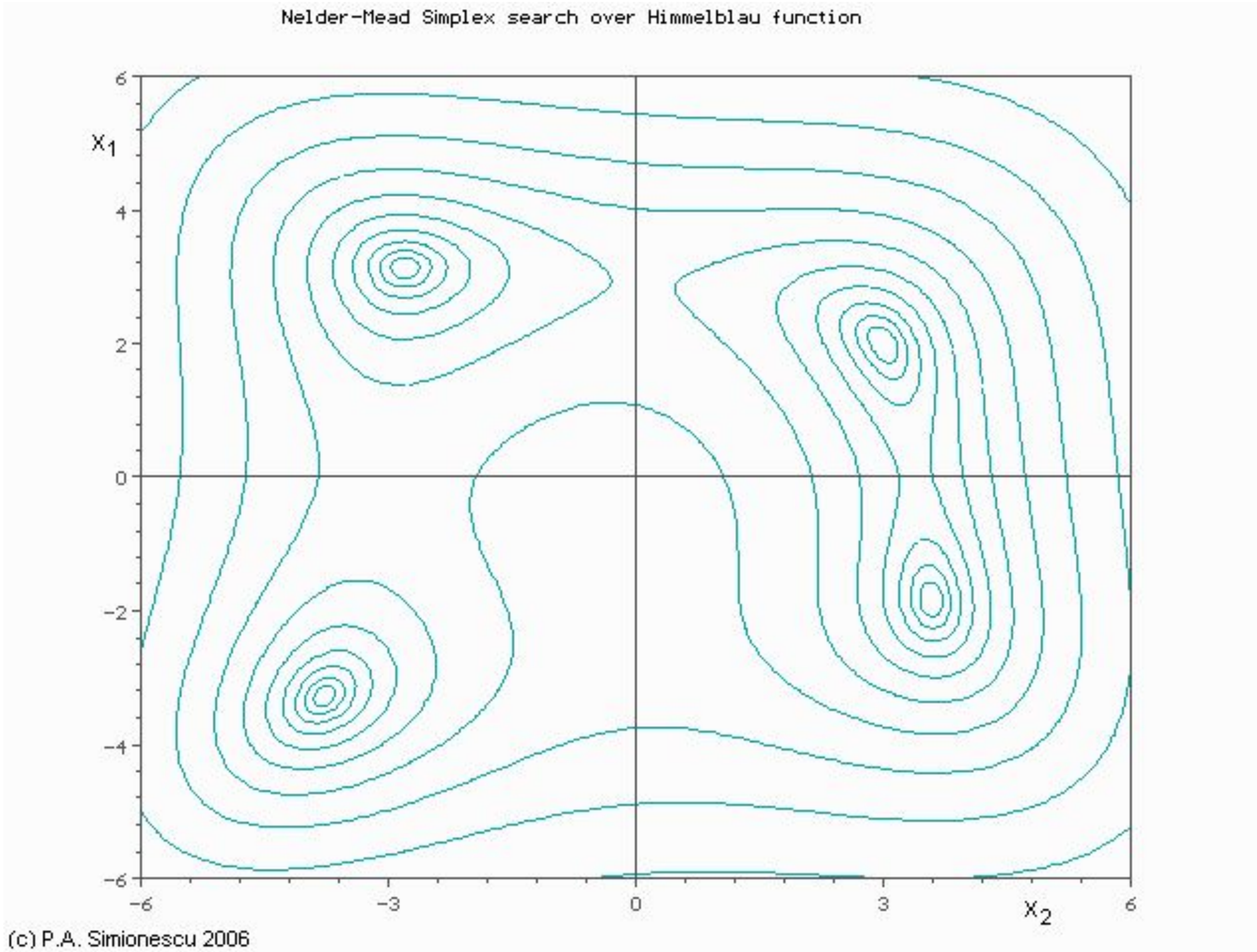


Таблица значений для 10 итераций

Best	Good	Worst
$f(0,1) = -8$	$f(1.0,0) = -5$	$f(0,0) = 0$
$f(1.5,1.5) = -15.75$	$f(0,1) = -8$	$f(1.0,0) = -5$
$f(0.25,3.75) = -20.187$	$f(1.5,1.5) = -15.75$	$f(0,1) = -8$
$f(0.25,3.75) = -20.187$	$f(1.75,4.25) = -20.1875$	$f(1.5,1.5) = -15.75$
$f(1.125,3.375) = -20.671$	$f(1.75,4.25) = -20.1875$	$f(0.25,3.75) = -20.1875$
$f(1.140,3.796) = -20.9638$	$f(1.125,3.375) = -20.6718$	$f(1.75,4.25) = -20.1875$
$f(1.140,3.796) = -20.9638$	$f(1.287,3.751) = -20.8668$	$f(1.125,3.375) = -20.6718$
$f(1.140,3.796) = -20.9638$	$f(1.236,3.874) = -20.9521$	$f(1.287,3.751) = -20.8668$
$f(0.990,4.002) = -20.9951$	$f(1.140,3.796) = -20.9638$	$f(1.2365,3.874) = -20.9520$
$f(0.990,4.002) = -20.9951$	$f(0.895,3.925) = -20.9855$	$f(1.140,3.796) = -20.9638$

Ответ:
После 10 итераций мы получаем достаточно точное приближение:
F(0.990, 4.002) = -20.9951
Аналитический экстремум функции достигается в F(1,4) = -21

Визуальный пример работы



Список литературы

1. КУРС «Многомерная оптимизация». Лекция 10. Метод Нелдера — Мида на сайте Института дистанционного обучения ИНТУИТ.
2. Метод Нелдера-Мида. Краткий алгоритм.
3. http://bigor.bmstu.ru/?cnt/?doc=MO/ch0606.mod/?cou=MO/base.cou
4. http://wp.wiki-wiki.ru/wp/index.php/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%9D%D0%B5%D0%BB%D0%B4%D0%B5%D1%80%D0%B0_%E2%80%94_%D0%9C%D0%B8%D0%B4%D0%B0
5. https://habr.com/ru/post/332092/