Лекция № 5

Методы нулевого порядка

(8) Метод Нелдера-Мида (деформируемого многогранника)

Алгоритм метода:

1) Задается начальная система точек (многогранник), включающая в себя $n+1$ точку: $X^{0(1)}, X^{0(2)} \dots X^{0(n+1)}$.

Замечание. Для функции 2-х переменных задается три начальные точки: $X^{0(1)}, X^{0(2)}, X^{0(3)}$.

2) Вычисляется значение функции во всех точках многогранника и выбирается:

- лучшая точка $X^{(л)}$: $f(X^{(л)}) = \min\limits_i \left[f(X^{k(i)} \right]$, здесь k − номер итерации, i − номер точки;
- худшая точка $X^{(х)}$: $f(X^{(х)}) = \max\limits_i \left[f(X^{k(i)} \right]$.

3) Далее заданная система из $n+1$ точки перестраивается, для этого строится центр тяжести системы заданных точек за исключением худшей:

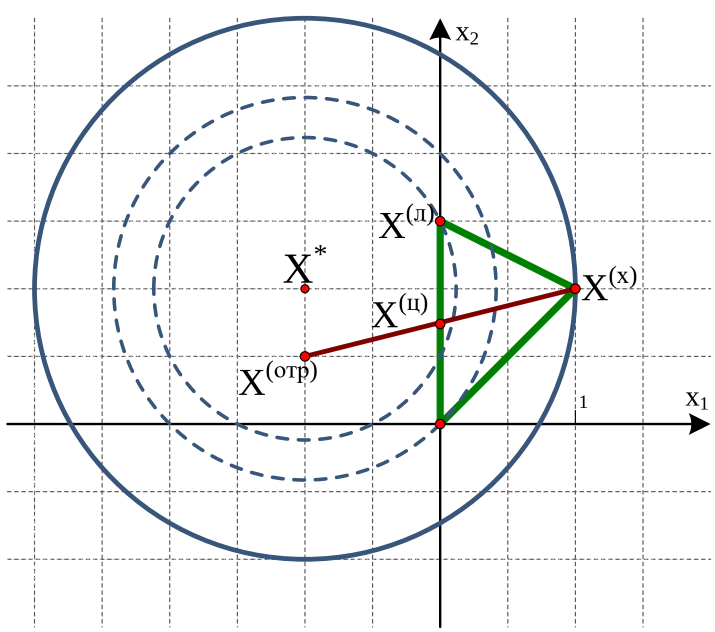
$$\boxed{X^{(ц)} = \frac{1}{n}\left(\sum_{i=1}^{n=1} X^{k(i)} - X^{(х)}\right)}$$

(для функции 2-х переменных точка $X^{(ц)}$ − середина отрезка, соединяющего точки за исключением худшей)

4) Выполняется операция ***отражение*** худшей точки через центр тяжести:

$$\boxed{X^{(отр)} = X^{(ц)} + \alpha \cdot (X^{(ц)} - X^{(х)})},$$

здесь $\alpha > 0$ − параметр отражения (рекомендуемое значение $\alpha = 1$).

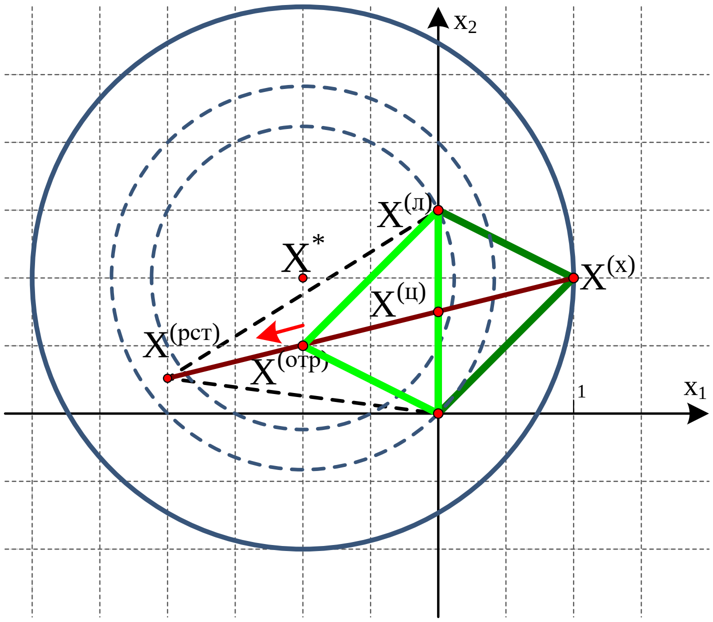


5) Формируется новая система точек (многогранник). Для этого в точке $X^{(отр)}$ вычисляется значение функции, полученное значение сравнивается с $f(X^{(л)})$:

- если $f(X^{(отр)}) < f(X^{(л)})$ – точка «лучше лучшей», то выполняется операция ***растяжение:***

  $\boxed{X^{(рст)} = X^{(ц)} + \gamma(X^{(отр)} - X^{(ц)})}$,

  здесь $\gamma > 0 \quad (\gamma \neq 0)$ – параметр растяжения (рекомендованное значение $\gamma \in [2, 3]$)



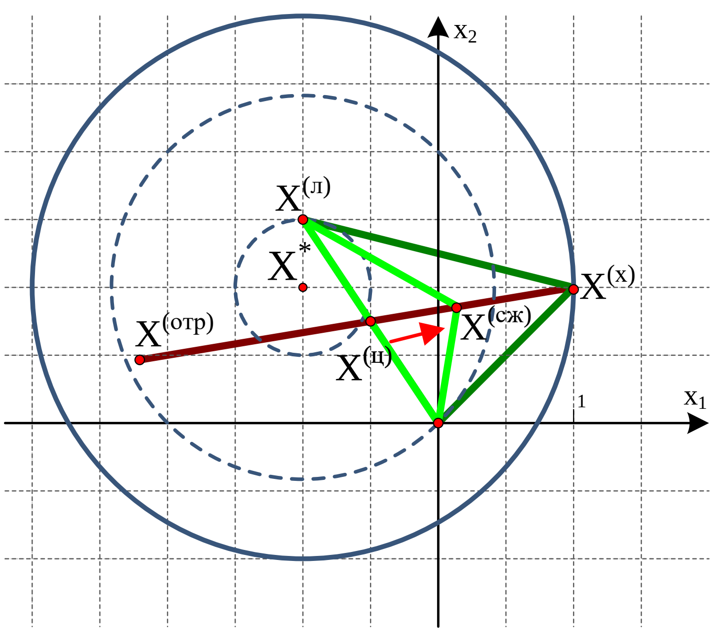
При этом если $f(X^{(рст)}) < f(X^{(отр)})$, то в новой системе точек точка $X^{(х)}$ будет заменена на $X^{(рст)}$, если же $f(X^{(рст)}) \geq f(X^{(отр)})$, то в новой системе точек точка $X^{(х)}$ будет заменена на $X^{(отр)}$.

На рисунке представлен случай, когда $f(X^{(рст)}) > f(X^{(отр)})$, точка $X^{(рст)}$ аннулируется, в новой системе точек $X^{(х)}$ будет заменена на $X^{(отр)}$.

- если $f(X^{(л)}) \leq f(X^{(отр)}) < f(X^{(х)})$ – точка «лучше худшей, но хуже лучшей», то выполняется операция ***сжатие:***

  $\boxed{X^{(сж)} = X^{(ц)} + \beta(X^{(х)} - X^{(ц)})}$,

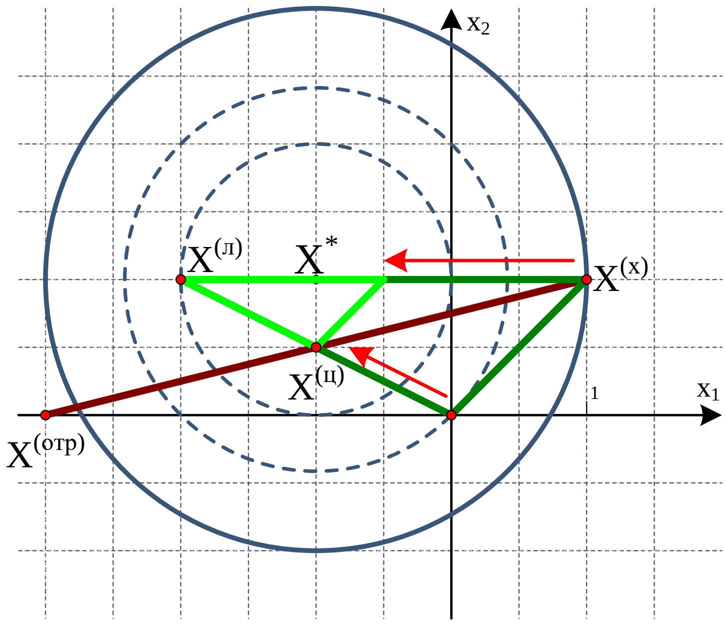
  здесь $\beta > 0 \quad (\beta \neq 0)$ – параметр сжатия (рекомендованное значение $\beta \in [0.4, 0.6]$).



При этом если $f(X^{(сж)}) < f(X^{(отр)})$, то в новой системе точек точка $X^{(х)}$ будет заменена на $X^{(сж)}$, если же $f(X^{(сж)}) \geq f(X^{(отр)})$, то в новой системе точек точка $X^{(х)}$ будет заменена на $X^{(отр)}$.

На рисунке представлен случай, когда $f(X^{(сж)}) < f(X^{(отр)})$, точка $X^{(х)}$ будет заменена на $X^{(сж)}$.

- если $f(X^{(отр)}) \ge f(X^{(х)})$ – точка «хуже худшей», то выполняется операция ***редукция***, при этом формируется новый многогранник, содержащий точку $X^{(л)}$ с уменьшенными вдвое сторонами:
  $\boxed{X^{k+1(i)} = X^{(л)} + 0,5(X^{k(i)} - X^{(л)}), \quad i = 1..n+1}$



Таким образом, в результате выполнения этого пункта алгоритма формируется новая система точек (многогранник), причем в случае возникновения операций растяжения и сжатия перестраивается только одна точка – $X^{(х)}$, в случае возникновения операции редукции – все точки, за исключением $X^{(л)}$.

6) Процедура 2)−5) повторяется до выполнения критерия окончания счета.

<u>Основной критерий окончания метода:</u> $\boxed{\sqrt{\frac{1}{n+1}\sum_{i=1}^{n+1}\left[f(X^{k(i)}) - f(X^{(ц)})\right]^2} \le \varepsilon}$.

<u>Дополнительные критерии окончания метода:</u>

- при выполнении предельного числа итераций: $k = M$;
- при однократном или двукратном одновременном выполнении двух условий:
  $\left\|X^{k+1} - X^{k}\right\| < \tilde{\varepsilon}, \qquad \left|f(X^{k+1}) - f(X^{k})\right| < \tilde{\varepsilon}$, где $\tilde{\varepsilon}$ – малое положительное число.

(9) Метод случайного поиска (адаптивный метод случайного спуска)

<u>Алгоритм метода</u>:

1) Задается начальная точка X^0 и начальное значение параметра r_0.

2) Строится система пробных точек (обычно 5−10 точек): $X^{пр(i)} = X^k + r_k \xi^i$, здесь k – номер итерации, ξ^k – случайный вектор единичной длины, i – номер пробной точки.

Построенные пробные точки оказываются лежащими на гиперсфере радиуса r_k (в случае двух переменных – на окружности радиуса r_k).

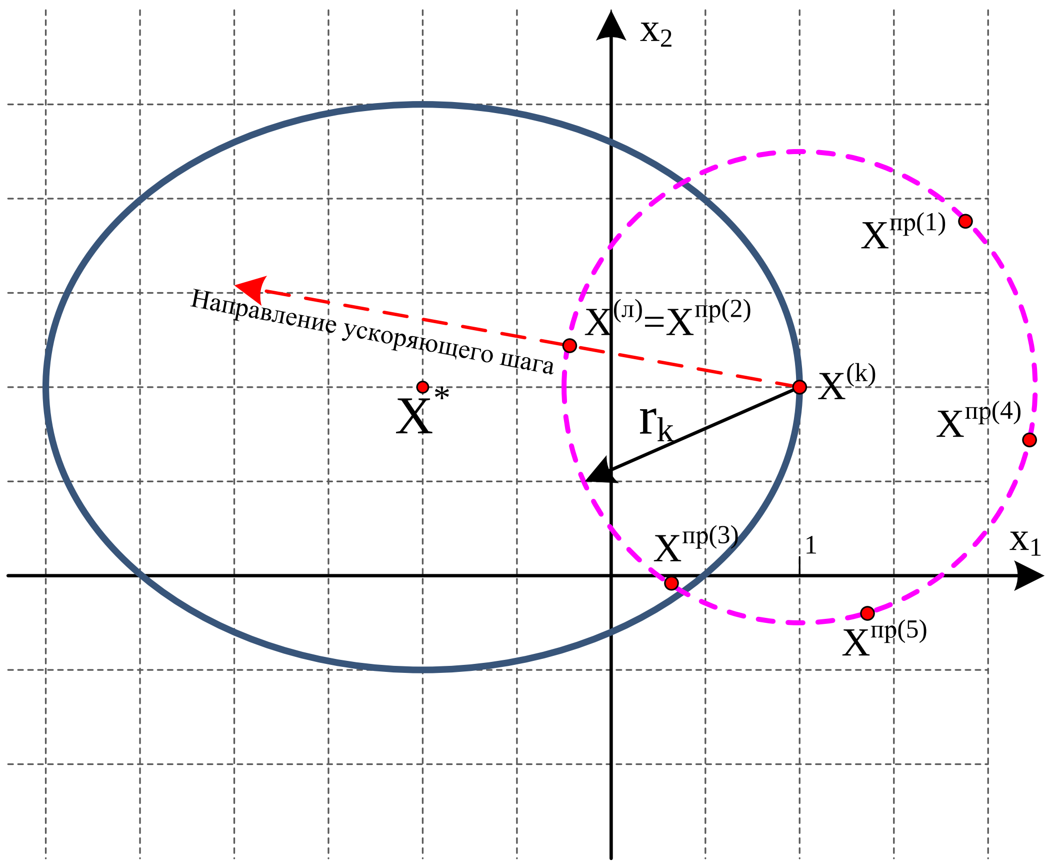
3) Для каждой пробной точки вычисляется значение функции $f(X^{пр(i)})$ и выбирается наилучшая $X^{(л)}$, для которой $f(X^{(л)}) = \min\left[f(X^{пр(i)})\right]$. Выбор может осуществляться как автоматически, так и при участии пользователя.

4) Проверяется условие: $f(X^{(л)}) < f(X^{k})$:

- если условие выполнено, то система пробных точек считается удачной, далее возможно два продолжения алгоритма:

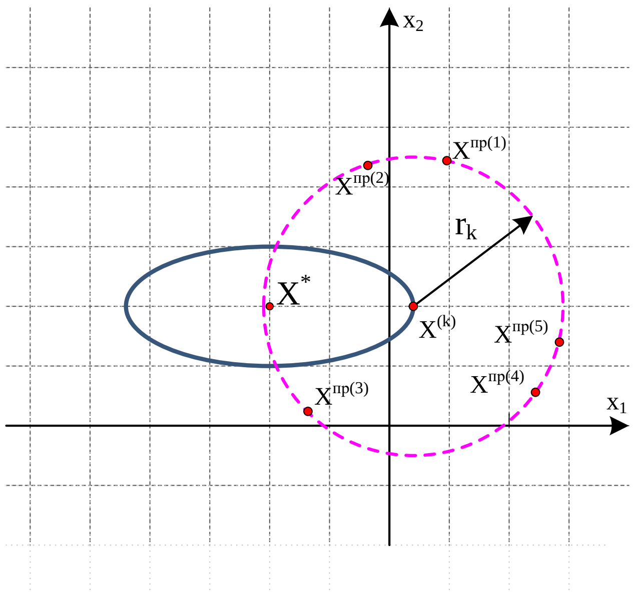
  4.1) $X^{k+1} = X^{(л)}$;

  4.2) в направлении, соединяющем точки X^{k} и $X^{(л)}$, делается ускоряющий шаг: $X^{k+1} = X^{(л)} + \lambda(X^{(л)} - X^{(k)})$, в этом случае, если оказывается, что $f(X^{k+1}) \geq f(X^{k})$, принимается $X^{k+1} = X^{(л)}$;

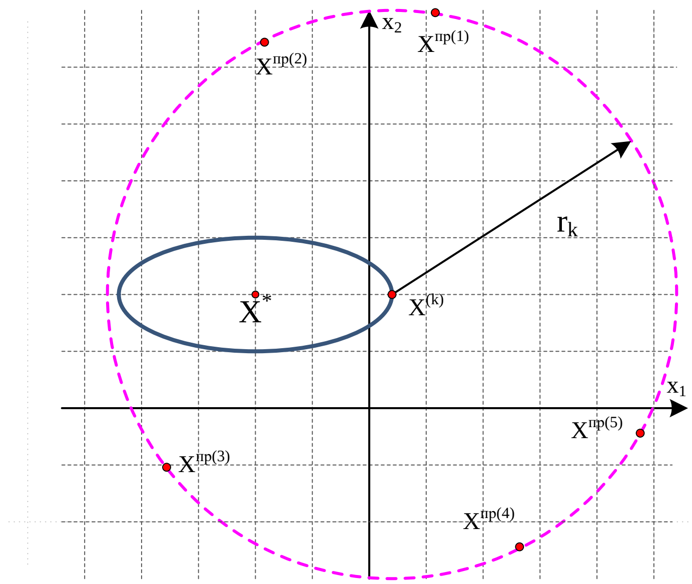


Удачная система пробных точек

- если условие не выполняется, делается попытка построить новую удачную систему пробных точек. если при этом число неудачных попыток достигает некоторого заданного числа P, текущий радиус r_k уменьшается (автоматически или пользователем).



Неудачная система пробных точек
(возможна повторная попытка)



Неудачная система пробных точек
(необходимо уменьшить радиус)

5) Процедура 2)−4) повторяется до выполнения критерия окончания счета.

Основной критерий окончания метода: $\boxed{r^k < \varepsilon}$.

Дополнительные критерии окончания метода:

- при выполнении предельного числа итераций: $k = M$;
- при однократном или двукратном одновременном выполнении двух условий: $\left\| X^{k+1} - X^k \right\| < \tilde{\varepsilon}$, $\left| f(X^{k+1}) - f(X^k) \right| < \tilde{\varepsilon}$, где $\tilde{\varepsilon}$ − малое положительное число.

(10) Метод конфигураций (Хука-Дживса)

Метод представляет собой комбинацию *исследующего (исследовательского) поиска* с циклическим изменением переменных и ускоряющего *поиска по образцу*.

Процесс поиска минимума функции всегда начинается с исследующего поиска.

Исследующий поиск осуществляется вдоль координатных направлений, результатом его являются так называемые точки базиса, в которых вычисляется значение функции $f(X)$.

Поиск по образцу осуществляется в направлении, соединяющем две последующие точки базиса. В точках полученных «по образцу» значение функции не вычисляется, они служат лишь для проведения в них исследующего поиска.

<u>Алгоритм метода</u>:

1) Задается начальная точка X^0 и начальные значение приращений $dx_1^0, dx_2^0 .... dx_n^0$. Точка X^0 называется точкой старого базиса.

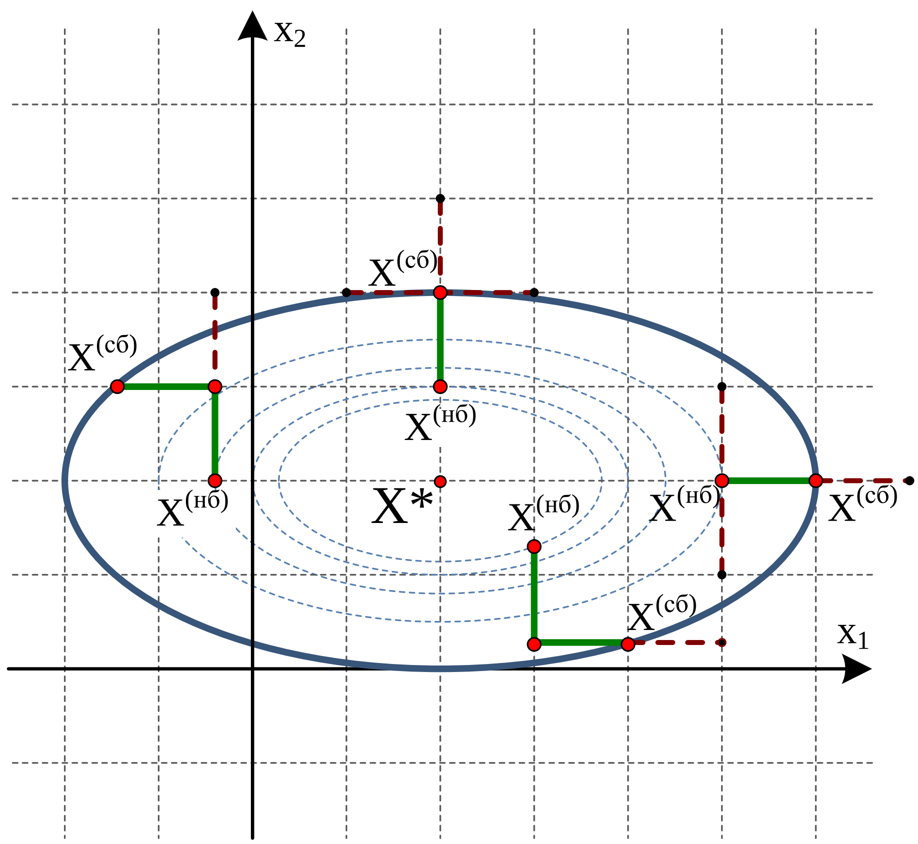
2) Проводится исследующий поиск, в результате которого каждая координата новой точки X^{k+1} вычисляется по алгоритму:

$$x_i^{k+1} = \begin{cases} x_i^k + dx_i^k, & \text{если } f(x_1^k, x_2^k ... x_i^k + dx_i^k ... x_n^k) < f(x_1^k, x_2^k ... x_i^k ... x_n^k) \\ x_i^k - dx_i^k, & \text{если } f(x_1^k, x_2^k ... x_i^k - dx_i^k ... x_n^k) < \min\left(f(x_1^k, x_2^k ... x_i^k ... x_n^k), f(x_1^k, x_2^k ... x_i^k + dx_i^k ... x_n^k)\right) \\ x_i^k, & \text{в противном случае} \end{cases}$$

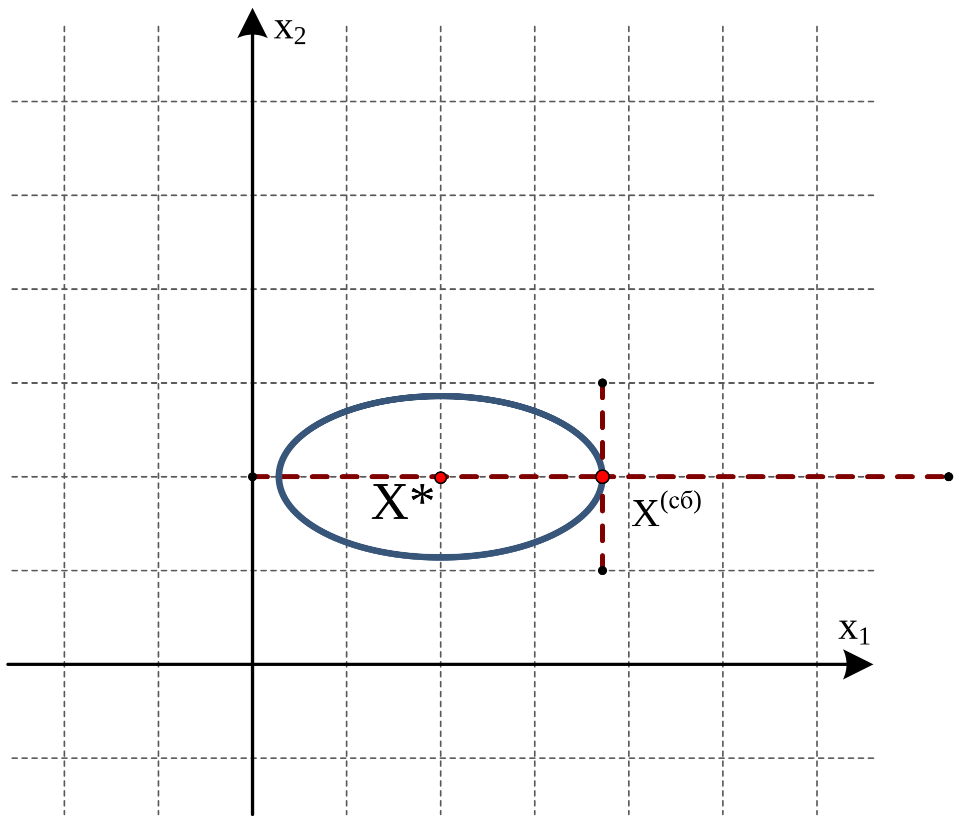
В результате исследующего поиска получается точка X^{k+1}.

Если при этом $X^{k+1} \neq X^k$, то X^{k+1} – точка нового базиса.

Если $X^{k+1} = X^k$, то исследующий поиск неудачен. В этом случае необходимо уменьшить значения приращений $dx_1^k, dx_2^k .... dx_n^k$ и повторить исследующий поиск.



Удачный исследующий поиск.
Последовательно проверяются направления «вправо-влево», затем «вверх-вниз», если функция возрастает, то точки аннулируются, если убывает – принимаются.
$X^{сб}$ – *точка старого базиса,* $X^{нб}$ – *точка нового базиса*



Неудачный исследующий поиск.

$X^{сб}$ – *точка старого базиса, точка нового базиса не может быть построена*

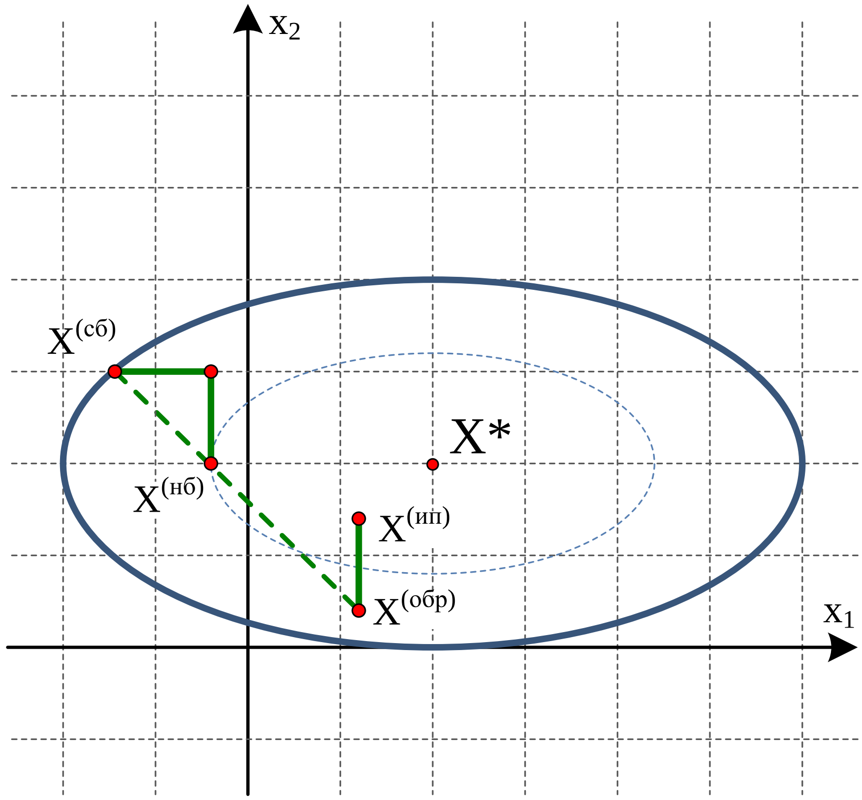
3) Из точки нового базиса может быть:

- продолжен исследующий поиск со старыми или новыми значениями приращений (шаг 2) алгоритма);
- проведен поиск по образцу по алгоритму: $X^{обр} = X^k + t_k(X^k - X^{k-1})$.

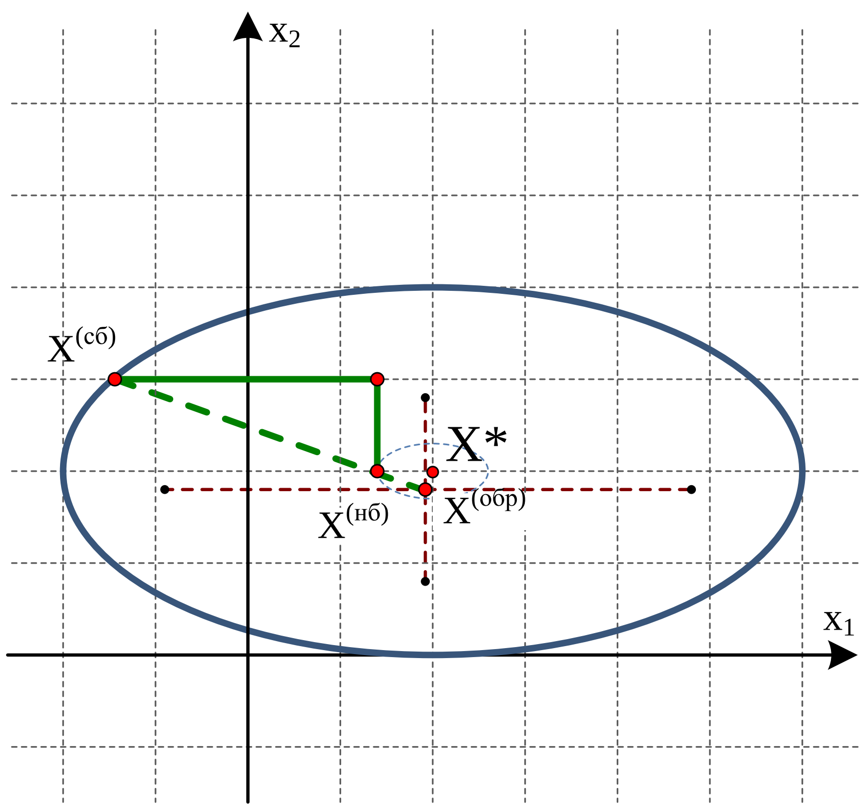
В точке $X^{обр}$ значение функции не вычисляется, из этой точки проводится исследующий поиск, в результате которого получается точка $X^{ип}$.

Если $X^{ип} \neq X^{обр}$, то точка $X^{k+1} = X^{ип}$ становится точкой нового базиса, а X^k – точкой старого базиса.

Если $X^{ип} = X^{обр}$, то поиск по образцу считается неудачным, точки $X^{ип}, X^{обр}$ – аннулируются, при этом точка X^k остается точкой нового базиса, а X^{k-1} – точкой старого базиса.



Удачный поиск по образцу



Неудачный поиск по образцу

4) Процедура 3) повторяется до выполнения критерия окончания счета.

Основной критерий окончания метода: $\boxed{dx_i^k < \varepsilon, \quad i = 1..n}$.

Дополнительные критерии окончания метода:

- при выполнении предельного числа итераций: $k = M$;
- при однократном или двукратном одновременном выполнении двух условий:
  $\left\| X^{k+1} - X^k \right\| < \tilde{\varepsilon}, \qquad \left| f(X^{k+1}) - f(X^k) \right| < \tilde{\varepsilon}$, где $\tilde{\varepsilon}$ − малое положительное число.