Нелинейное программирование

Задача безусловной оптимизации

$$f(x) \to \min \qquad x \in X, \qquad X \in R^n$$

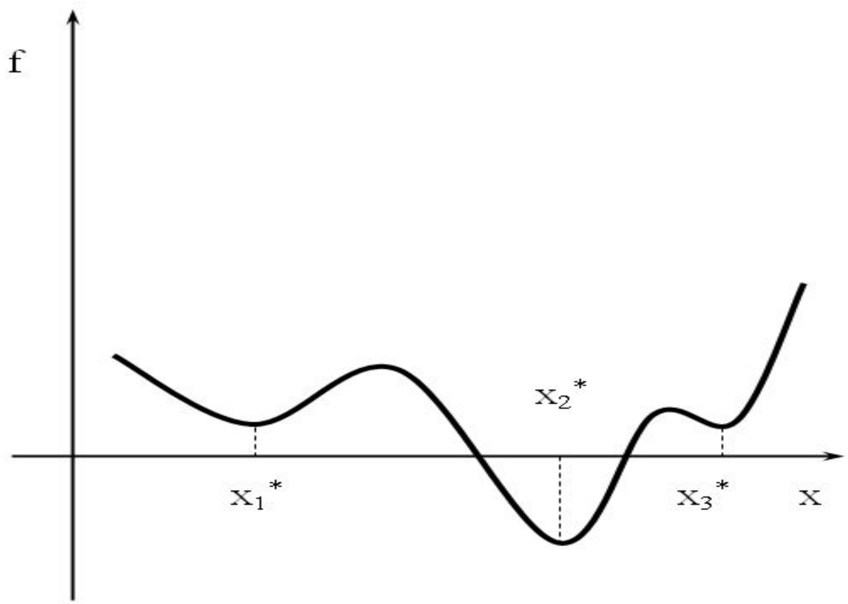
Минимум x^*:

1. Глобальный - $x^* \in X$

$$f(x^*) \le f(x) \text{ для всех } x \in X$$

2. Локальный - $x^* \in X$

$$f(x^*) \le f(x) \text{ для всех } x \in X \bigcap U_\varepsilon(x^*)$$
$$U_\varepsilon(x^*) = \{x \in R^n \mid \|x - x^*\| \le \varepsilon\}$$



Условия оптимальности

1. Необходимое условие минимума

а) функция одной переменной

$$f'(x^*) = 0$$

б) функция многих переменных

градиент $f'(\bar{x}^*) = \left(\dfrac{\partial f(\bar{x}^*)}{\partial x_1}, \ldots, \dfrac{\partial f(\bar{x}^*)}{\partial x_n} \right) = 0$

$\dfrac{\partial f(\bar{x}^*)}{\partial x_i} = 0 \quad i = \overline{1, n} \qquad \bar{x}^*$ - стационарная точка

2. Достаточное условие минимума

а) функция одной переменной

$$f''(x^*) > 0$$

б) функция многих переменных

гессиан $G_{ij} = \dfrac{\partial^2 f(\bar{x}^*)}{\partial x_i \partial x_j} \quad i, j = \overline{1, n} \quad G_{ii} = \dfrac{\partial^2 f(\bar{x}^*)}{\partial x_i^{\,2}} > 0$

Одномерная задача безусловной оптимизации

Существование решения – теорема Вейерштрасса:

целевая функция $f(x)$ непрерывна на множестве σ $([a,b])$ - x_1 и x_2 такие

$$f(x) \geq f(x_1) \quad - \ x_1 \text{ минимум}$$
$$f(x) \leq f(x_2) \quad - \ x_2 \text{ максимум}$$

Классический подход: $f(x)$ кусочно-непрерывна и кусочно-гладкая для $x \in \sigma$

минимум в $x \in \sigma$:

а) $f'(x) = 0$

б) $f'(x)$ - не существует

в) $f(x)$ терпит разрыв

г) x^* - граничная для σ $(x^* = a$ и $x^* = b)$

Численные методы – **оптимальные** – получение решения задачи минимизации ЦФ с требуемой точностью при вычислении только значений функции $f(x_i)$ в меньшем количестве промежуточных точек x_i

Общая схема численных методов

1. специально подобранные n точек $x_1, x_2, \ldots, x_n$ $x_i \in [a, b]$

расчет $f(x_i)$ $i = \overline{1, n}$

2. выбор $\bar{x}_n$ из условия $f(\bar{x}_n) = \min\limits_{1 \le i \le n} f(x_i)$

3. определение из значений $a, b, x_1, x_2, \ldots, x_n$

$$a_n \leftarrow \bar{x}_n \rightarrow b_n$$

ближайшая

слева *справа*

4. x^* - это $x_n^* \in [a_n, b_n]$

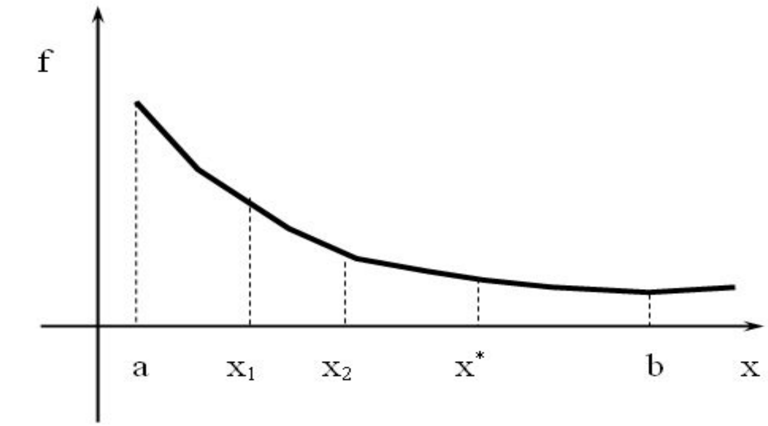
а) $x_n^* = \bar{x}_n$

б) $x_n^* = \dfrac{1}{2}(a_n + b_n)$

$x^* \in [a_n, b_n]$ - для унимодальных функций

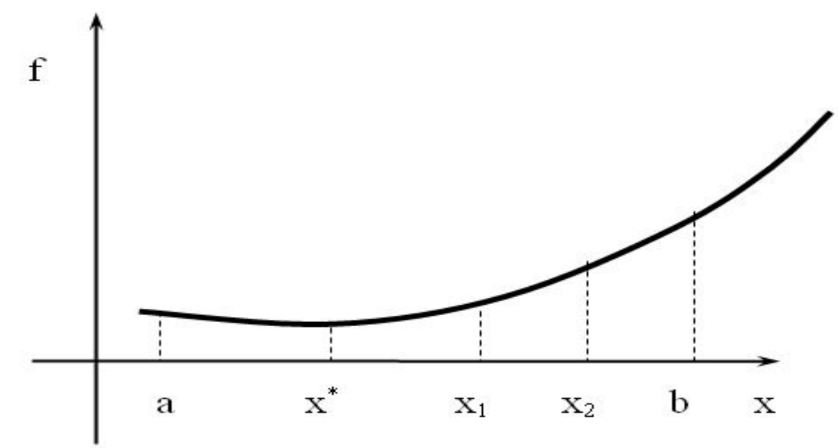
Унимодальные функции

$f(x) \quad x \in \sigma \quad ([a,b])$ существует $x^* \in \sigma$



$f(x_1) > f(x_2) \quad x_1 < x_2 < x^* \quad x_1, x_2 \in \sigma$

$\left[a, x^*\right]$ - $f(x)$ *убывает*



$f(x_1) < f(x_2) \quad x^* < x_1 < x_2 \quad x_1, x_2 \in \sigma$

$\left[x^*, b\right]$ – $f(x)$ *возрастает*

<u>Лемма</u>: $f(x)$ унимодальная на σ

$x_1, x_2 \in \sigma \qquad x_1 < x_2$

а) $f(x_1) \le f(x_2) \quad x^* \le x_2 \quad x^* \in [a, x_2]$

б) $f(x_1) \ge f(x_2) \quad x^* \ge x_1 \quad x^* \in [x_1, b]$

<u>Интервал неопределенности (локализации)</u> x^*

$\Delta_n(f)$ - <u>длина интервала локализации</u> x^* после вычисления значений $f(x)$ в точках $x_1, x_2, \ldots, x_n \in [a, b]$

Оптимальный поиск – способ построения последовательности $x_1, x_2, \ldots, x_n$, при которой $\Delta_n(f) \to \min$

Способ построения последовательности

$$x_1, x_2, \ldots x_n$$

Пассивный поиск
Метод перебора

а) $f(x)$ - не унимодальная

n - число участков разбиения

$$h = \frac{b-a}{n} \qquad x_k = a + kh \;\; k = \overline{0,n}$$

$$y_k = f(x_k)$$

$$m_n = \min_{0 \le k \le n} y_k$$

$$\lim_{n \to \infty} m_n \to \min \;\; \to \text{ выбор } n?$$

оценка погрешности?

б) $f(x)$ - унимодальная

$$y_i^* = \min_{0 \le k \le n} y_k \qquad x^* \in [x_{i-1}, x_{i+1}]$$

если $|x_{i+1} - x_{i-1}| > \varepsilon$, новое разбиение $[x_{i-1}, x_{i+1}]$

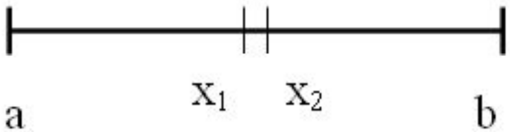
Последовательный поиск

1. метод дихотомии
2. метод золотого сечения
3. метод чисел Фибоначчи

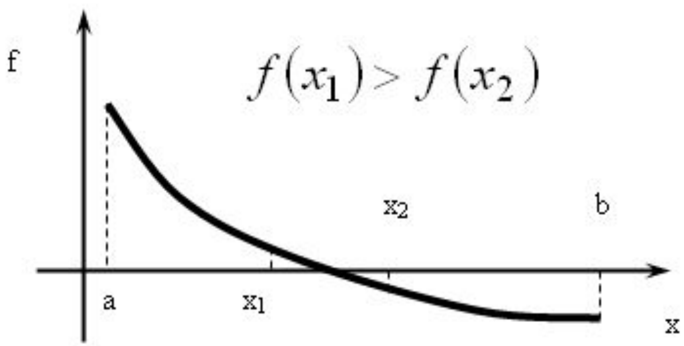
Отличие методов – выбор узлов x_i для сужения интервала локализации

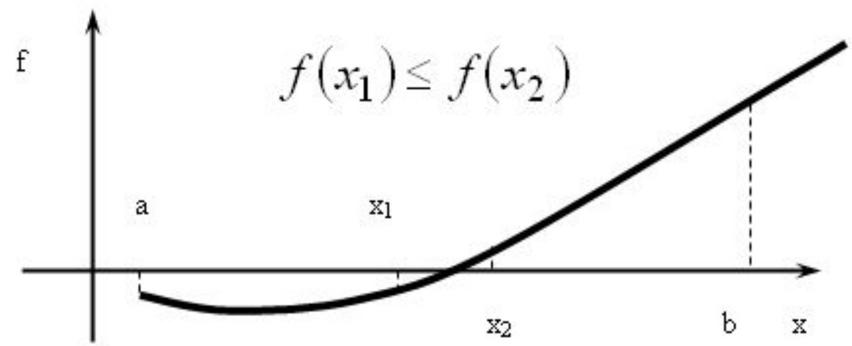
Метод дихотомии

параметр метода $\varepsilon = const \quad 0 < \varepsilon < (b-a)$



$$x_1 = \frac{a+b-\varepsilon}{2} \quad x_2 = \frac{a+b+\varepsilon}{2} = a+b-x_1$$





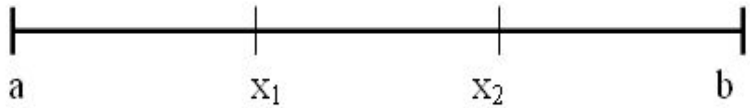
$$[a,b] \to [x_1,b] = [a_1,b_1] \qquad [a,b] \to [a,x_2] = [a_1,b_1]$$

$$x^* \in [a_1,b_1] \quad b_1 - a_1 = \frac{b-a}{2} + \frac{\varepsilon}{2} = \frac{b-a-\varepsilon}{2} + \varepsilon$$

шаг l: $\quad [a_l, b_l] \quad b_l - a_l = \dfrac{b-a-\varepsilon}{2^l} + \varepsilon \;\; (b_l - a_l) \leq \delta \;\; \delta > \varepsilon \quad l \geq \log_2\left(\dfrac{b-a-\varepsilon}{\delta - \varepsilon}\right)$

погрешность вычисления минимума $\quad \delta^*_{МД} \leq |b_l - a_l| = (b-a-\varepsilon)/2^l + \varepsilon \quad l = n/2$

Метод золотого сечения



$$\frac{b-a}{b-x_1}=\frac{b-x_1}{x_1-a}=1{,}618 \qquad \frac{b-a}{x_2-a}=\frac{x_2-a}{b-x_2}=1{,}618=\frac{1+\sqrt{5}}{2}$$

$$x_1=a+\frac{3-\sqrt{5}}{2}(b-a)=a+0{,}382(b-a) \quad x_2=a+\frac{\sqrt{5}-1}{2}(b-a)=a+0{,}618(b-a)=a+b-x_1$$

x_1 - золотое сечение $[a,b]$ и $[a,x_2]$ $\qquad$ x_2 - золотое сечение $[a,b]$ и $[x_1,b]$

<u>Шаг 1.</u> $a=a_1$ $b=b_1$ расчет x_1 и x_2 $\quad f(x_1)$ и $f(x_2)$ $\rightarrow$ $[a_2,b_2] \Leftrightarrow \begin{cases}[a_1,x_2]\\ [x_1,b_1]\end{cases}$

$$b_2-a_2=x_2-a=b-x_1=\frac{\sqrt{5}-1}{2}(b-a) \quad \bar{x}_2\in[a_2,b_2] \quad f(\bar{x}_2)=\min_{1\le i\le 2} f(x_i)$$

$\bar{x}_2$ - золотое сечение $[a_2,b_2]$

<u>Шаг 2.</u> $x_3=a_2+b_2-\bar{x}_2$ расчет $f(x_3)\rightarrow[a_3,b_3]$ $\quad b_3-a_3=\left(\left(\sqrt{5}-1\right)/2\right)^2(b-a)$

$\bar{x}_3\in[a_3,b_3] \quad f(\bar{x}_3)=\min_{1\le i\le 3} f(x_i)$ $\qquad$ $\bar{x}_3$ - золотое сечение $[a_3,b_3]$

Метод золотого сечения (продолжение)

Шаг $(n-1)$. $\bar{x}_{n-1}$ - золотое сечение $[a_{n-1}, b_{n-1}]$ $f(\bar{x}_{n-1}) = \min\limits_{1 \le i \le (n-1)} f(x_i)$

расчет $x_n = a_{n-1} + b_{n-1} - \bar{x}_{n-1}$ и $f(x_n) \rightarrow [a_n, b_n]$

$$b_n - a_n = \left(\frac{\sqrt{5}-1}{2}\right)^{n-1} (b-a)$$

$$\bar{x}_n \in [a_n, b_n] \qquad f(\bar{x}_n) = \min_{1 \le i \le n} f(x_i)$$

$\bar{x}_n$ - золотое сечение $[a_n, b_n]$

$$\bar{x}_n = a_n + \frac{3-\sqrt{5}}{2}(b_n - a_n) \quad \text{или} \quad \bar{x}_n = a_n + \frac{\sqrt{5}-1}{2}(b_n - a_n)$$

$$\delta^*_{МЗС} \le \max\left\{\begin{matrix} |\bar{x}_n - a_n| \\ |b_n - \bar{x}_n| \end{matrix}\right\} = \left(\frac{\sqrt{5}-1}{2}\right)^{n} (b-a)$$

Сравнение МД $(\varepsilon \to 0)$ и МЗС - одинаково n (количество вычислений $f(x_i)$ $i = \overline{1, n}$)

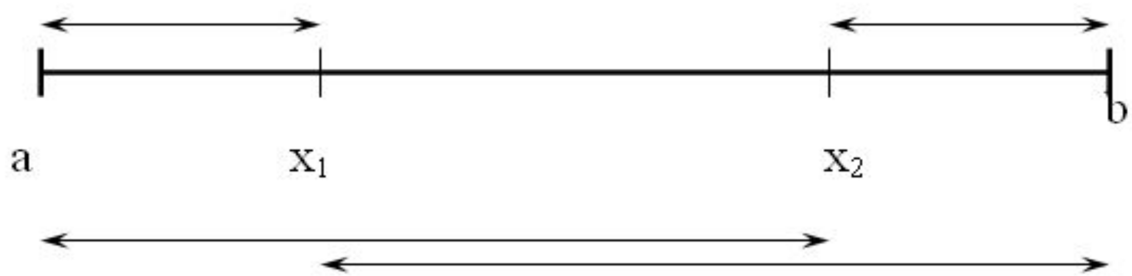
$$\frac{\delta^*_{МЗС}}{\delta^*_{МД}} = \left(\left(\frac{\sqrt{5}-1}{2}\right)^{n} (b-a)\right) \Bigg/ \frac{b-a}{2^l} = \left\{ l = \frac{n}{2} \right\} = \left(\frac{\sqrt{2}}{2}\left(\sqrt{5}-1\right)\right)^{n} \approx (0{,}874)^n$$

Метод чисел Фибоначчи

$$F_1 = 1 \qquad F_2 = 1 \qquad F_{n+2} = F_{n+1} + F_n \quad n = 1, 2, \ldots$$

Рекуррентная формула $\qquad F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$

$$\frac{F_n}{F_{n+2}}(b-a) \qquad\qquad \frac{F_n}{F_{n+2}}(b-a)$$



$$\frac{F_{n+1}}{F_{n+2}}(b-a)$$

Пусть $n \geq 2$ (n - количество вычислений значений функции $f(x)$)

$$a_1 = a \qquad b_1 = b$$

Метод чисел Фибоначчи (продолжение)

<u>Шаг 1.</u> расчет $x_1 = a + \frac{F_n}{F_{n+2}}(b-a)$ и $x_2 = a + \frac{F_{n+1}}{F_{n+2}}(b-a) = a + b - x_1$

$$f(x_1) \quad \text{и} \quad f(x_2) \;\rightarrow\; [a_2, b_2] \Leftrightarrow \begin{cases} [a, x_2] \\ [x_1, b] \end{cases}$$

$$b_2 - a_2 = \frac{F_{n+1}}{F_{n+2}}(b-a) \quad \bar{x}_2 \in [a_2, b_2] \quad f(\bar{x}_2) = \min_{1 \le i \le 2} f(x_i)$$

$$\bar{x}_2 = x_2' = a_2 + \frac{F_n}{F_{n+2}}(b-a) = a_2 + \frac{F_n}{F_{n+1}}(b_2 - a_2)$$

или

$$\bar{x}_2 = x_2'' = a_2 + \frac{F_{n-1}}{F_{n+2}}(b-a) = a_2 + \frac{F_{n-1}}{F_{n+1}}(b_2 - a_2)$$

Метод чисел Фибоначчи (продолжение)

<u>Шаг 2.</u> расчет $x_3 = a_2 + b_2 - \bar{x}_2$ и $f(x_3) \rightarrow [a_3, b_3]$

$$b_3 - a_3 = \frac{F_n}{F_{n+1}}(b_2 - a_2) = \frac{F_n}{F_{n+2}}(b - a) \quad \bar{x}_3 \in [a_3, b_3] \quad f(\bar{x}_3) = \min_{1 \le i \le 3} f(x_i)$$

$$\bar{x}_3 = x_3' = a_3 + \frac{F_{n-1}}{F_n}(b_3 - a_3) = a_3 + \frac{F_{n-1}}{F_{n+2}}(b - a)$$

или

$$\bar{x}_3 = x_3'' = a_3 + \frac{F_{n-2}}{F_n}(b_3 - a_3) = a_3 + \frac{F_{n-2}}{F_{n+2}}(b - a)$$

<u>Шаг</u> $(n-1)$ расчет $x_n = a_{n-1} + b_{n-1} - \bar{x}_{n-1}$ и $f(x_n) \rightarrow [a_n, b_n]$

$$b_n - a_n = \frac{F_3}{F_4}(b_{n-1} - a_{n-1}) = \frac{F_3}{F_{n+2}}(b - a) \quad \bar{x}_n \in [a_n, b_n] \quad f(\bar{x}_n) = \min_{1 \le i \le n} f(x_i)$$

$$\bar{x}_n = x_n' = a_n + \frac{F_1}{F_3}(b_n - a_n) = a_n + \frac{F_1}{F_{n+2}}(b - a)$$

или

$$\bar{x}_n = x_n'' = a_n + \frac{F_2}{F_3}(b_n - a_n) = a_n + \frac{F_2}{F_{n+2}}(b - a)$$

Метод чисел Фибоначчи (продолжение)

Погрешность вычисления минимума

$$\delta^*_{МЧФ} = \max\begin{Bmatrix} |\bar{x}_n - a_n| \\ |b_n - \bar{x}_n| \end{Bmatrix} = \frac{F_2}{F_{n+2}}(b-a) = \frac{F_1}{F_{n+2}}(b-a) = \frac{b-a}{F_{n+2}}$$

Сравнение МЧФ и МЗС $(n \to \infty)$ $\qquad F_n \approx \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$

$$\frac{\delta^*_{МЗС}}{\delta^*_{МЧФ}} = {(b-a)\left(\frac{\sqrt{5}-1}{2}\right)^n} \Bigg/ {\frac{b-a}{\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+2}}} = \left(\frac{\sqrt{5}+1}{2}\right)^2 \frac{1}{\sqrt{5}} = 1{,}1708$$

<u>Сравнение интервальных методов</u>

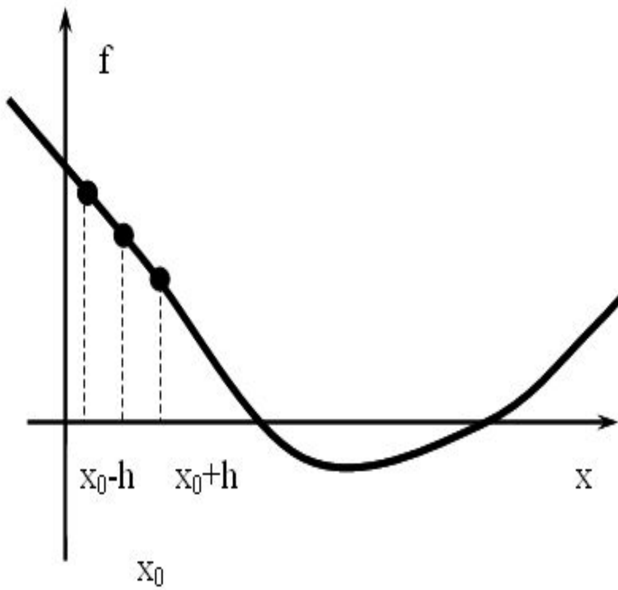
<u>МД</u>: 1. простой в реализации 2. применим для поиска минимума разрывной функции
3. большой объем вычислений для обеспечения заданной точности δ

<u>МЗС</u>: 1. предел для чисел Фибоначчи $\lim\limits_{n\to\infty} F_n / F_{n+1} = \left(\sqrt{5}-1\right)/2$

2. расположение x_i не зависит от n

<u>МЧФ</u>: оптимальный метод – при том же n наименьший интервал локализации x^*

Метод последовательного перебора



Скорость сходимости зависит от:

1. x_0
2. h

Для ускорения спуска – локальные свойства функции вблизи x_0

а) $f'(x_0)$ - скорость и направление убывания функции

б) $f''(x_0)$ - направление выпуклости функции

Дано: $x_0 \quad h \quad \varepsilon$

x_0 и $x_0 + h$

а) $f(x_0) < f(x_0 + h) \rightarrow h = -h \quad x_0 + h$

б) $f(x_0 - h) = f(x_0) = f(x_0 + h) \rightarrow h = h/4$

при $|h| \le \varepsilon \qquad x_{\min} = x_0$

в) $f(x_0) > f(x_0 + h) \rightarrow$ спуск до

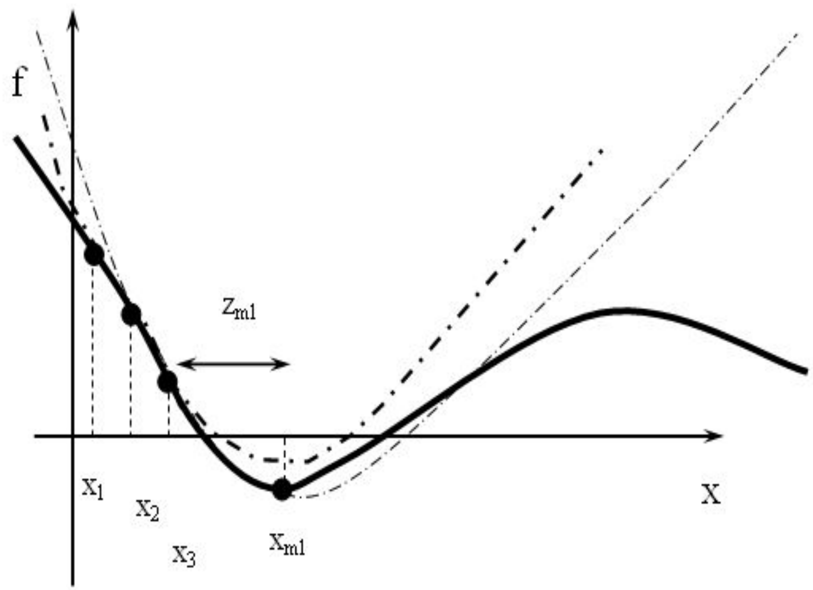
$f(x_0 + kh) > f(x_0 + (k-1)h)$

$$x_{\min} \in [x_0 + (k-2)h, x_0 + kh]$$

при $|h| \le \varepsilon \quad x_{\min} = x_0 + (k-1)h$

при $|h| > \varepsilon \quad h = -h/4$ - повторение процедуры спуска от $x_0 + kh$

Метод квадратичной параболы (Пауэлла)



Вблизи локального минимума

$$\begin{cases} f'(x^*) = 0 \\ f''(x^*) > 0 \end{cases}$$

Аппроксимация квадратичной параболой

Заданы: $x_0 \quad h \quad \varepsilon$

I) выбор начальных приближений

$$f(x_0) > f(x_0 + h) \qquad \begin{aligned} x_1 &= x_0 \\ x_2 &= x_0 + h \\ x_3 &= x_0 + 2h \end{aligned}$$

$$f(x_0) < f(x_0 + h) \qquad \begin{aligned} x_1 &= x_0 - h \\ x_2 &= x_0 \\ x_3 &= x_0 + h \end{aligned}$$

расчет $y_1 = f(x_1)$ $y_2 = f(x_2)$ $y_3 = f(x_3)$

II) Построение квадратичной параболы

$$\varphi(x - x_3) = a(x - x_3)^2 + b(x - x_3) + c = az^2 + bz + c$$

$$z = x - x_3 \quad z_1 = x_1 - x_3 \quad z_2 = x_2 - x_3$$

Система уравнений $\begin{cases} y_3 = c \\ y_2 = az_2^2 + bz_2 + c \\ y_1 = az_1^2 + bz_1 + c \end{cases}$

$$a = \frac{(y_1 - y_3)z_2 - (y_2 - y_3)z_1}{z_1 z_2 (z_1 - z_2)}$$

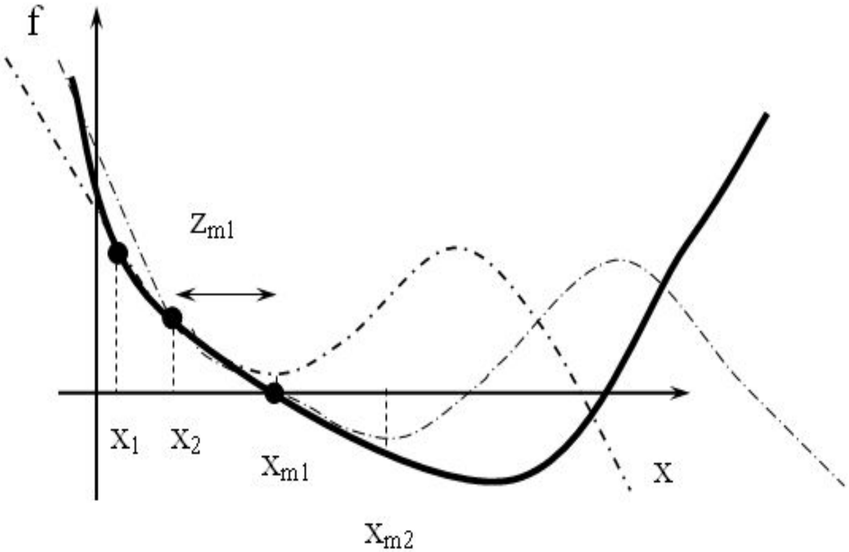
$$b = \frac{(y_1 - y_3){z_2}^2 - (y_2 - y_3){z_1}^2}{z_1 z_2 (z_2 - z_1)}$$

При $a > 0$ $\min \varphi(x - x_3)$ $\quad z_{m1} = -b/2a \quad x_{m1} = x_3 + z_{m1}$

1. при $|z_{m1}| < \varepsilon \rightarrow x_{\min} = x_{m1}$ 2. при $|z_{m1}| > \varepsilon \rightarrow$

$$x_1 = x_2 \;\; x_2 = x_3 \;\; x_3 = x_{m1} \rightarrow \text{II)}$$

Метод кубической параболы (Давидона)



I) Выбор начальных приближений

$$x_1 = x_0 \qquad x_2 = \begin{cases} x_0 + h & \left(f(x_0+h) < f(x_0)\right) \\ x_0 - h & \left(f(x_0-h) < f(x_0)\right) \end{cases}$$

расчет $\quad y_1 = f(x_1) \qquad y_2 = f(x_2)$

$\qquad D_1 = f'(x_1) \qquad D_2 = f'(x_2)$

II) Построение кубической параболы

$$\varphi(x - x_2) = a(x-x_2)^3 + b(x-x_2)^2 + c(x-x_2) + d =$$
$$= az^3 + bz^2 + cz + d$$

$z = x - x_2 \quad z_1 = x_1 - x_2$ проверка условия $\dfrac{D_2 - D_1}{h} > 0$

Система уравнений

$$\begin{cases} \varphi(x_1) = y_1 \\ \varphi(x_2) = y_2 \\ \varphi'(x_1) = D_1 \\ \varphi'(x_2) = D_2 \end{cases} \leftrightarrow \begin{cases} az_1^3 + bz_1^2 + cz_1 + d = y_1 \\ d = y_2 \\ 3az_1^2 + 2bz_1^2 + c = D_1 \\ c = D_2 \end{cases}$$

$$a = \frac{D_1 - D_2 - 2(y_1 - y_2 - D_2 z_1)/z_1}{z_1^2}$$

$$b = \frac{D_2 - D_1 - 3(y_1 - y_2 - D_2 z_1)/z_1}{z_1}$$

Минимум кубической параболы $\varphi'(z) = 0$

$$3az^2 + 2bz + c = 0 \rightarrow z_{m1} = \frac{-2b \pm \sqrt{4b^2 - 12ac}}{6a}$$

$x_{m1} = x_2 + z_{m1}$ 1. при $|z_{m1}| < \varepsilon \rightarrow x_{\min} = x_{m1}$

2. при $|z_{m1}| > \varepsilon \rightarrow x_1 = x_2 \;\; x_2 = x_3 \;\; x_3 = x_{m1} \rightarrow$ II)

при $b^2 - 3ac < 0$ - точка перегиба $z_{m1} = -b/3a$