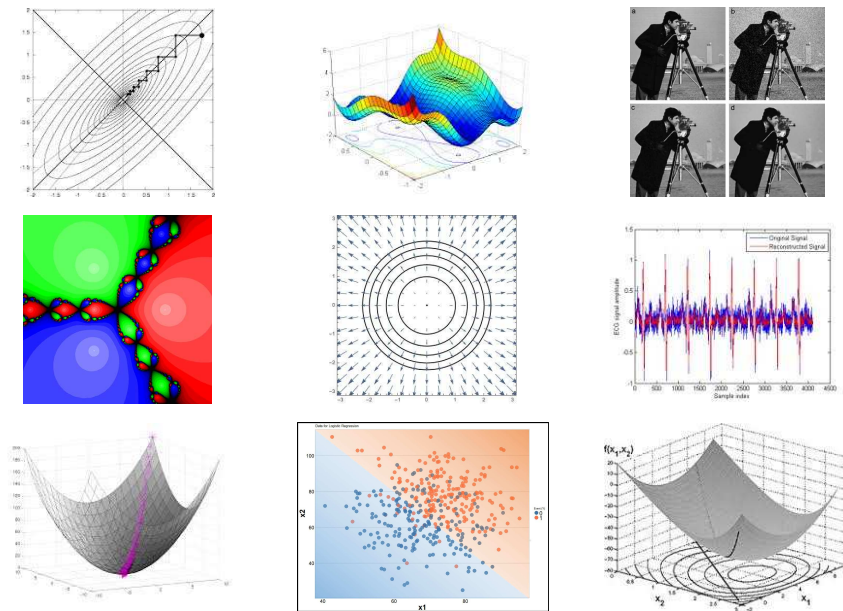
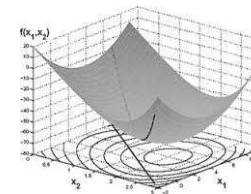
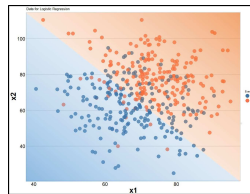
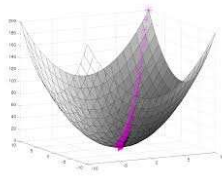
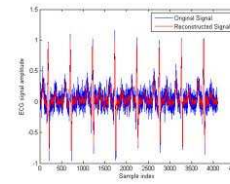
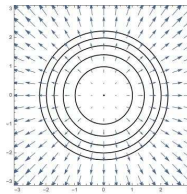
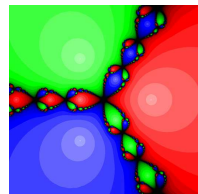
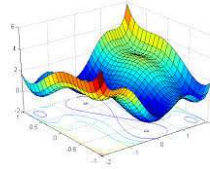
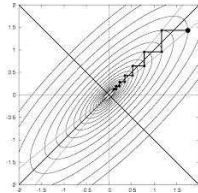




Metódy voľnej optimalizácie - letný semester 2018

M. Trnovská, KAMŠ, FMFI UK



Metódy minimalizácie funkcie jednej premennej

Motivácia

$$\min_{x \in \mathbb{R}^n} f(x) \quad (U1)$$

Úloha (U1) sa zvyčajne rieši nejakým **iteračným algoritmom**, ktorý generuje postupnosť bodov

$$x^0, x^1, x^2, \dots \in \mathbb{R}^n, \quad f(x^k) \rightarrow \hat{x}, \text{ pre } k \rightarrow \infty$$

Vo väčšine algoritmov sa používa iteračná schéma

$$x^{k+1} = x^k + \lambda_k s^k.$$

Metódy s optimálnym krokom:

$$\lambda_k = \arg \min_{\lambda} \varphi(\lambda) = \arg \min_{\lambda} f(x^k + \lambda s^k)$$

Formulácia úlohy

$$\min_{x \in \mathbb{R}} f(x) \quad (1)$$

Riešenie úlohy prebieha v dvoch fázach:

I. **Separovanie minima** - úloha (1) sa prevedie na úlohu

$$\min_{a \leq x \leq b} f(x) \quad (2)$$

II. **Riešenie úlohy (2)**

Budeme predpokladať, že úloha (2) je riešiteľná. To nastane za predpokladu, že f je **unimodálna** na $[a, b]$.

Definícia:

Funkciu $f(x)$ definovanú na intervale $I \subset \mathbb{R}$ nazývame **unimodálnou**, ak existuje bod $x_0 \in I$ tak, že pre každú dvojicu bodov x_1, x_2 z intervalu I takú, že

$$x_2 < x_1 < x_0,$$

alebo

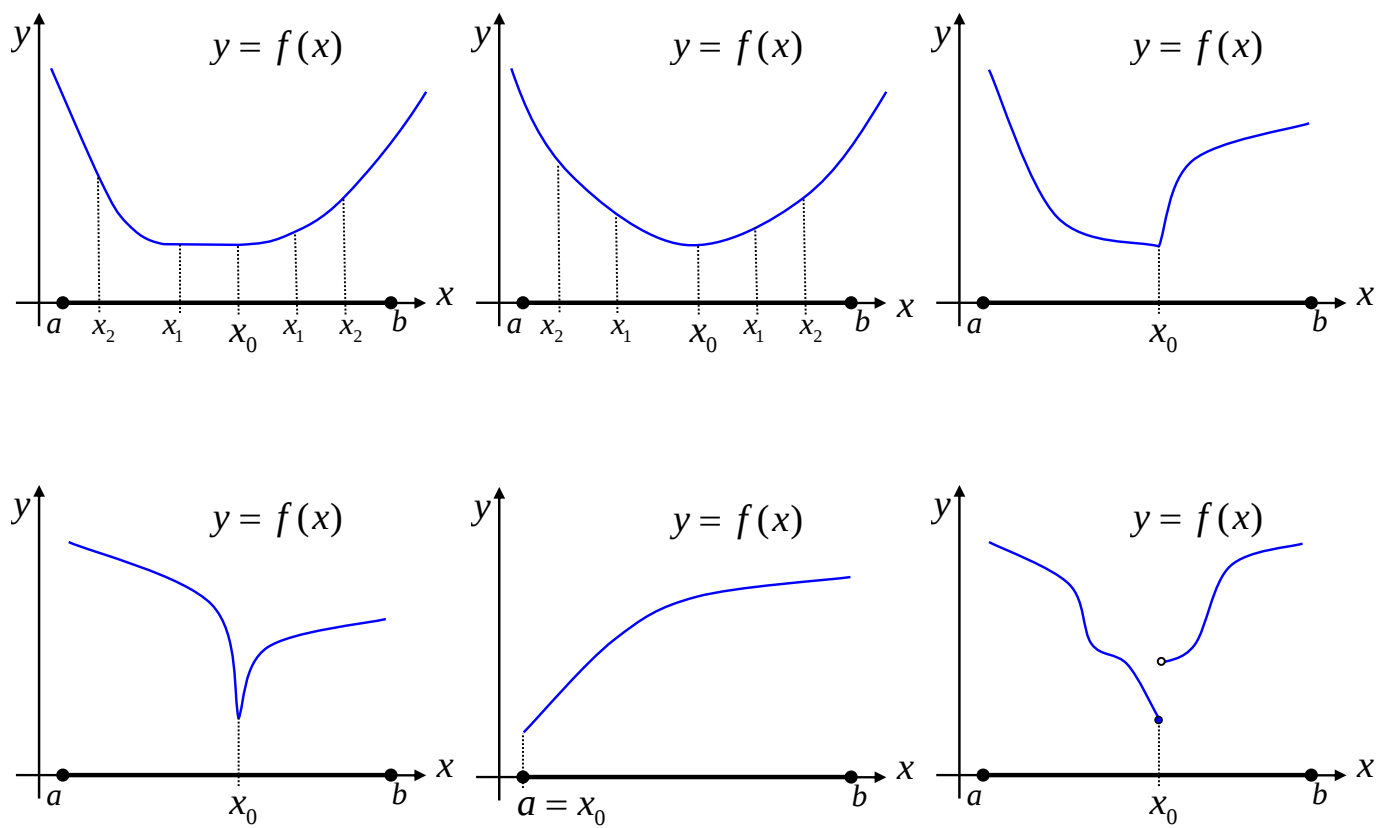
$$x_0 < x_1 < x_2,$$

platí

$$f(x_0) \leq f(x_1) \leq f(x_2). \quad (\star)$$

Unimodálnu funkciu $f(x)$ nazývame **rýdzo unimodálnou**, ak vo vzťahu $(\star)$ platia ostré nerovnosti.

Zrejme bod x_0 je bodom minima funkcie $f(x)$ vzhľadom na I .



Klasifikácia metód riešenia

Z hľadiska riešenia úlohy (2):

- **Metódy intervalovej aproximácie**
 - Hľadá sa dostatočne jemný interval neurčitosti obsahujúci bod minima x_0 .
 - Založené na bezprostrednom porovnávaní funkčných hodnôt .
 - Rýchlosť konvergenzie nezávisí od tvaru funkcie - vo všeobecnosti pomalá.
- **Metódy bodovej aproximácie (Interpolačné metódy)**
 - Problém určenia bodu $\bar{x}$ z dostatočne malého okolia bodu x_0 .
 - Založené na interpolácii funkcie f nejakou vhodnou funkciou.
 - Rýchlosť konvergenzie závisí od stupňa zhody funkcie f s interpolačnou funkciou.

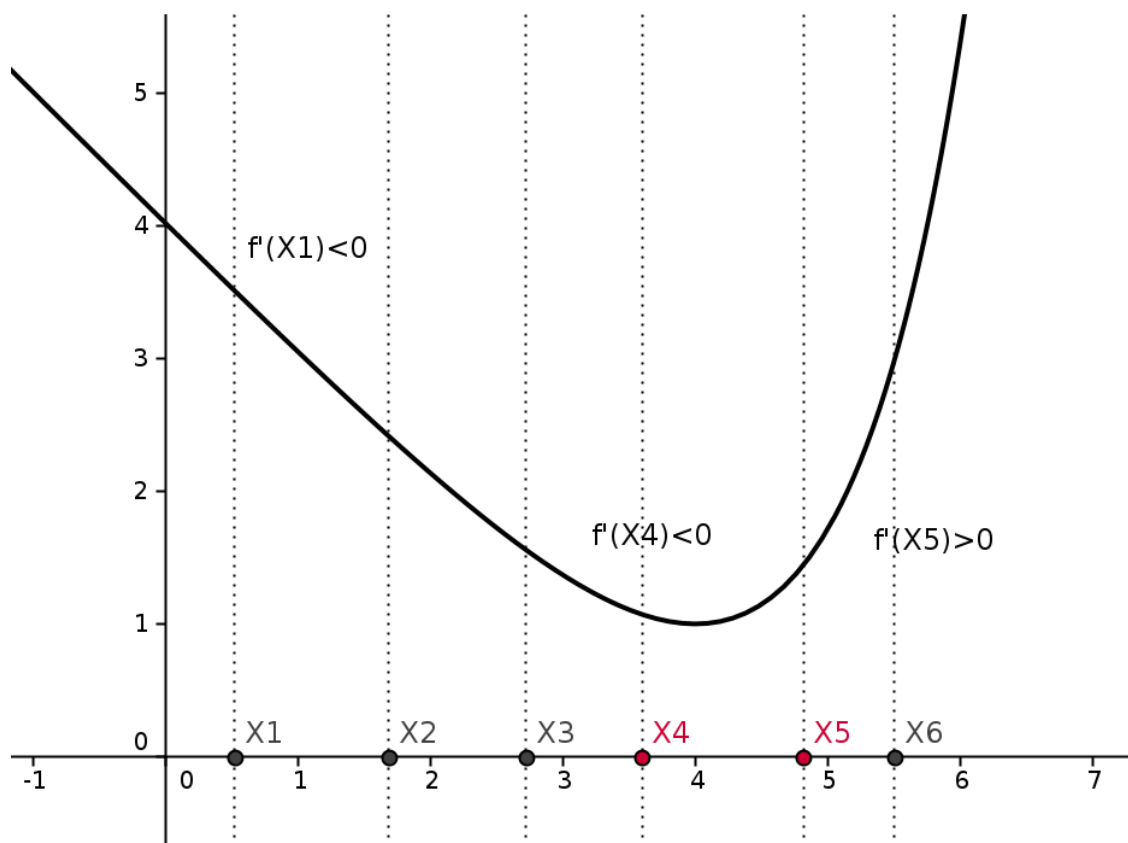
Klasifikácia metód riešenia

Z hľadiska použitej informácie:

- **Metódy nultého rádu**
 - Používajú len funkčné hodnoty
 - Pomalá konvergencia
- **Metódy prvého rádu**
 - Používajú aj hodnoty prvej derivácie
 - Lepšia konvergencia
- **Metódy druhého rádu**
 - Používajú aj hodnoty druhej derivácie
 - Najlepšia konvergencia, ale veľký objem výpočtov

Separovanie minima

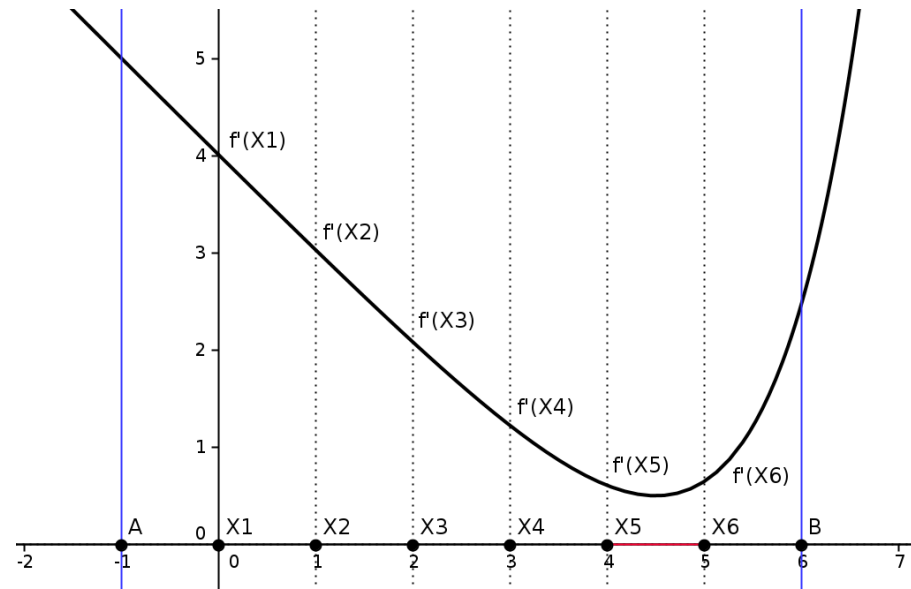
- Metóda 1. rádu
- Metóda 0. rádu



Metódy intervalovej aproximácie

- **Metódy simultánnych experimentov**
 - ak treba vopred rozhodnúť o polohe vyhodnocovaných bodov
 - ekvidistančná sieť alebo m -ekvidistančná sieť
 - metóda 1. rádu, metóda 0. rádu
- **Metódy postupných experimentov**
 - Metóda bisekcie (1. rádu)
 - Dichotomická metóda (0.rádu)
 - Metóda Fibonacci
 - Metóda zlatého rezu

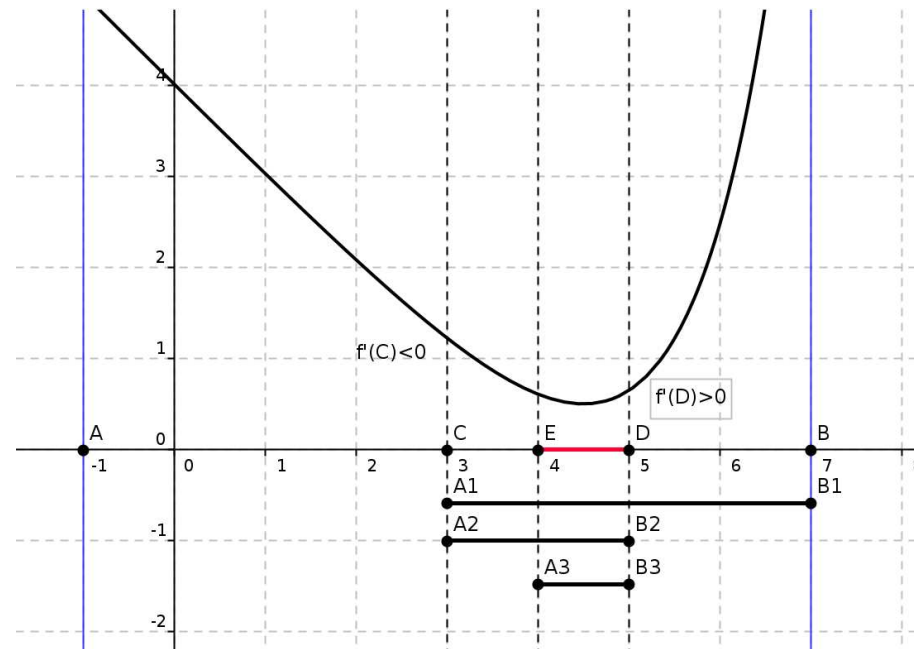
Metódy simultánných experimentov



Na obrázku je metóda simultánných experimentov 1. rádu. Platí $f'(X_5) < 0$ a $f'(X_6) > 0$. Interval neurčitosti obsahujúci minimum funkcie f je teda $[X_5, X_6]$.

Metódy postupných experimentov

Metóda bisekcie



Po k -tej iterácii sa interval zredukuje na dĺžku $B_k - A_k = \frac{B-A}{2^k}$.

Dichotomická metóda - analogická metóda 0. rádu

Metóda Fibonacci

Daný je interval $[a, b]$ a počet povolených experimentov (výpočtov) n .

Využíva sa Fibonacciho postupnosť

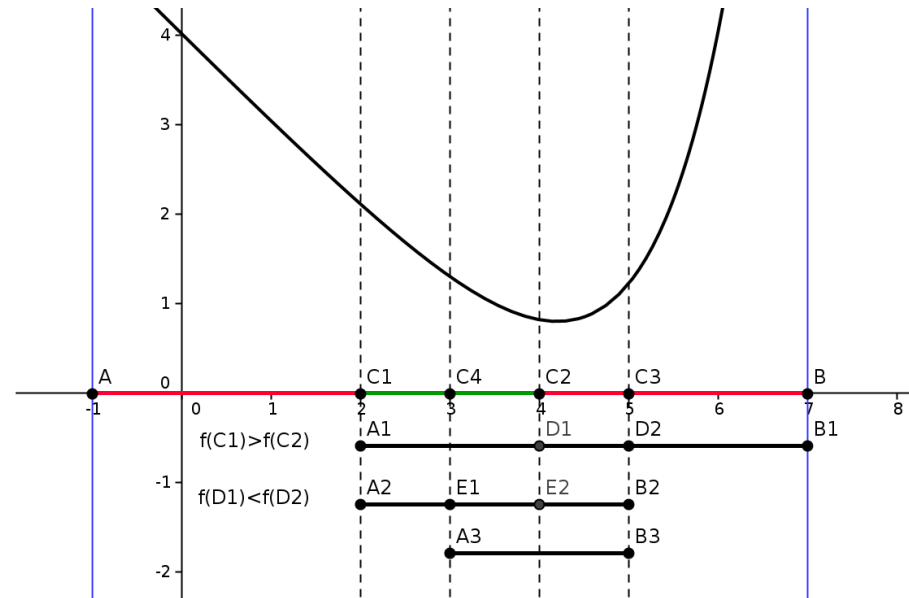
$$F_0 = F_1 = 1, F_{n+1} = F_n + F_{n-1}$$

n	0	1	2	3	4	5	6	7	8	9	10	11
F_n	1	1	2	3	5	8	13	21	34	55	89	144

Interval $[a, b]$ sa rozdelí na F_n dielikov.

V prvej iterácii sa vypočíta funkčná hodnota v bodoch

$$c_1 = a + \frac{F_{n-2}}{F_n}(b - a), \quad c_2 = a + \frac{F_{n-1}}{F_n}(b - a).$$



Fibonacciho metóda s povoleným počtom experimentov $n = 5$, interval $[a, b]$ sa delí na $F_n = F_5 = 8$ dielikov.

Metóda zlatého rezu

Nemá nedostatky Fibonacciho metódy.

Interval $[a, b]$ sa delí symetricky podľa pravidla zlatého rezu.

Faktor redukcie v každej iterácii je $\frac{1}{\varphi} = \frac{\sqrt{5}-1}{2}$.

Algoritmus zlatého rezu

Vstupy: $f(x), a, b, \varepsilon$

Schéma algoritmu:

1. Vypočítame $\phi, z_1 = 2 - \phi, z_2 = \phi - 1$.
2. Uzly $c_1 < c_2$ volíme podľa pravidla zlatého rezu:

$$c_1 = a + z_1(b - a), \quad c_2 = a + z_2(b - a).$$

3. Ak $f(c_1) < f(c_2)$

$$b := c_2, \quad c_2 := c_1, \quad c_1 := a + z_1(b - a)$$

inak (ak $f(c_1) \geq f(c_2)$)

$$a := c_1, \quad c_1 := c_2, \quad c_2 = a + z_2(b - a).$$

Metódy bodovej aproximácie (interpolačné metódy)

Vstup:

interpolačné uzly $x_1, x_2, \dots, x_k$;

hodnoty $f_i = f(x_i)$, $f'_i = f'(x_i)$, $f''_i = f''(x_i)$ pre nejaké i ;

daná trieda funkcií závisiaca od $r \geq k$ parametrov.

Schéma algoritmu:

1. Konštrukcia interpolačnej funkcie $\phi(x)$ - založená na výpočte hodnôt príslušných parametrov tak, aby v interpolačných uzloch platilo $\phi(x_i) = f_i$, resp. $\phi'(x_i) = f'_i$, $\phi''(x_i) = f''_i$.
2. Výpočet minima $\bar{x}$ interpolačnej funkcie $\phi(x)$, ktoré aproximuje hľadané minimum $\hat{x}$ funkcie $f(x)$.
3. Ak $\bar{x}$ je “dostatočne dobrou” aproximáciou hľadaného minima $\hat{x}$, tak končíme. V opačnom prípade bodom $\bar{x}$ nahradíme “najhorší” z použitých interpolačných uzlov. Vypočítame $\bar{f} = f(\bar{x})$, prípadne aj $\bar{f}' = f'(\bar{x})$ a vraciame sa na krok 1.

Kvadratická interpolácia minima

Interpolačná funkcia: $\phi(x) = a(x - z)^2 + b(x - z) + c$

Jej minimum: $\bar{x} = z - \frac{b}{2a}$

Prípad jedného interpolačného uzla

Funkcia $f(x)$ sa v okolí uzla x_1 aproximuje Taylorovým polynómom II. stupňa:

$$f(x) \approx f(x_1) + f'(x_1)(x - x_1) + \frac{1}{2}f''(x_1)(x - x_1)^2.$$

Newtonova metóda: $x_{k+1} = x_k - \frac{f'_k}{f''_k}$

Príklad.

Minimalizácia funkcie

$$f(x) = -5x^5 + 4x^4 - 12x^3 + 11x^2 - 2x$$

na intervale $[-\frac{1}{2}, \frac{1}{2}]$

Aplikovaním Newtonovej metódy so štartovacím bodom $x_0 = -1/3$ a toleranciou $\varepsilon = 10^{-11}$ (pre hodnotu derivácie) dostávame v jednotlivých iteráciách

$x = -0.3333333333333333$

$x = -0.074697173620458$

$x = 0.064413437535524$

$x = 0.105955490584666$

$x = 0.109826339072980$

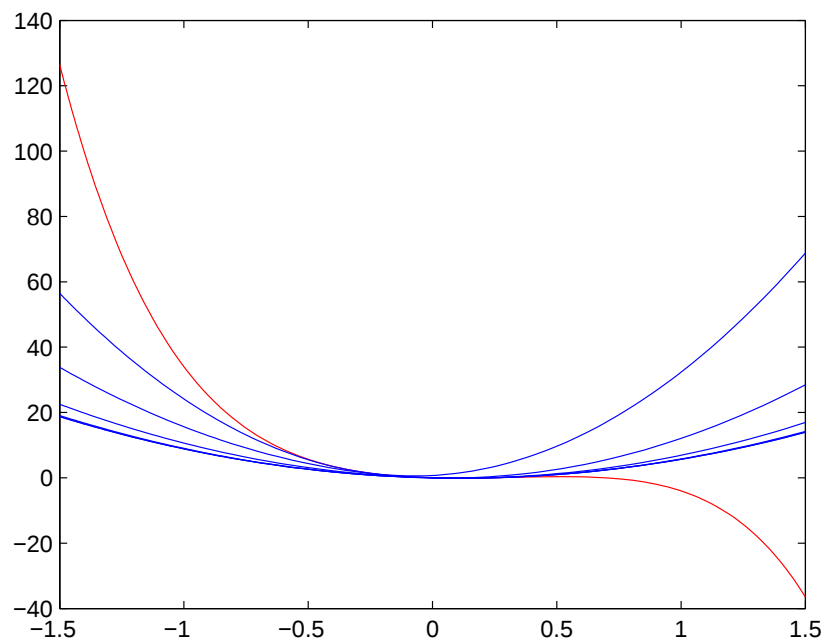
$x = 0.109859912568469$

$x = 0.109859915091411$

Kvadratické aproximácie funkcie

$$f(x) = -5x^5 + 4x^4 - 12x^3 + 11x^2 - 2x$$

v jednotlivých iteráciách:



Kvadratická interpolácia - prípad dvoch interpolačných uzlov

1. Interpolačná informácia: $f'_1 < 0 < f'_2$:

- Zrejme bod aproximácie minima $\bar{x}$ sa nachádza medzi uzlami x_1 a x_2 .
- Je vhodné zvoliť parameter $z = \frac{x_1+x_2}{2}$.
- Definuje sa pomocná veličina $h = \frac{x_2-x_1}{2}$.

Interpolačné podmienky:

$$\phi'(x_i) = 2a(x_i - z) + b = f'_i, \quad i = 1, 2.$$

Formula pre aproximáciu minima:

$$\bar{x} = \frac{1}{2} \left[(x_1 + x_2) - (x_2 - x_1) \frac{f'_1 + f'_2}{f'_2 - f'_1} \right] = x_1 - (x_2 - x_1) \frac{f'_1}{f'_2 - f'_1}$$

Príklad:

Minimalizácia funkcie $f(x) = -5x^5 + 4x^4 - 12x^3 + 11x^2 - 2x$

Metóda konverguje pomalšie ako Newtonova metóda:

1	-0.008413461538	
2	0.203427968074	
3	0.062553206455	
4	0.091712386562	
5	0.103055070795	
6	0.107333877600	
7	0.108925972245	
8	0.109515136929	23 0.109859914982
9	0.109732707592	24 0.109859915051
10	0.109812991159	25 0.109859915076
11	0.109842607272	26 0.109859915086
12	0.109853531312	27 0.109859915089
13	0.109857560536	28 0.109859915091
14	0.109859046655	29 0.109859915091

Kvadratická interpolácia - prípad dvoch interpolačných uzlov

1. Interpolačná informácia: $f'_1 < 0, f_1, f_2$

- Opäť je vhodné zvoliť parameter $z = \frac{x_1 + x_2}{2}$.
- Definujme pomocné veličiny

$$h = \frac{x_2 - x_1}{2}, \quad f_{12} = \frac{f_2 - f_1}{x_2 - x_1}.$$

Interpolačné podmienky:

$$\phi(x_i) = a(x_i - z)^2 + b(x_i - z) + c = f_i, \quad i = 1, 2$$

$$\phi'(x_1) = 2a(x_1 - z) + b = f'_1.$$

Formula pre aproximáciu minima:

$$\bar{x} = \frac{1}{2} \left[(x_1 + x_2) - (x_2 - x_1) \frac{f_{12}}{f_{12} - f_1'} \right]$$

resp.

$$\bar{x} = x_1 - (x_2 - x_1) \frac{f_1'}{2(f_{12} - f_1')}$$

Kvadratická interpolácia - prípad troch interpolačných uzlov

$$x_1 < x_2 < x_3, f_1 > f_2 < f_3, z = \frac{x_1 + x_3}{2}, f_{ij} = \frac{f_j - f_i}{x_j - x_i}$$

Interpolačné podmienky:

$$\phi(x_i) = a(x_i - z)^2 + b(x_i - z) + c = f_i, \quad i = 1, 2, 3$$

Interpolačná formula je

$$\hat{x} = \frac{1}{2} \left(x_1 + x_3 - \frac{f_3 - f_1}{f_{23} - f_{12}} \right)$$

Interpolácia minima kvadratickým splajnom

Interpoláčná funkcia:

$$\phi(x) = \begin{cases} \psi_1(x) = a_1(x - z)^2 + b(x - z) + c & x \leq z \\ \psi_2(x) = a_2(x - z)^2 + b(x - z) + c & x \geq z \end{cases}$$

kde z je bod zlepenia kvadratických parabol $\psi_1(x), \psi_2(x)$ a $a_1, a_2 > 0$.

$$\begin{aligned} \phi(z) &= \psi_1(z) = \psi_2(z) = c, \\ \phi'(z) &= \psi_1'(z) = \psi_2'(z) = b. \end{aligned}$$

Bod minima funkcie $\phi(x)$ je

$$\bar{x} = \begin{cases} z - \frac{b}{2a_1} & b \geq 0 \\ z - \frac{b}{2a_2} & b \leq 0 \end{cases}$$

Vstupná informácia: uzly $x_1 < x_2$, hodnoty $f_1, f_2, f'_1 < 0 < f'_2$

Interpolačné podmienky:

$$\begin{aligned}a_1(x_1 - z)^2 + b(x_1 - z) + c &= f_1 \\a_2(x_2 - z)^2 + b(x_2 - z) + c &= f_2 \\2a_1(x_1 - z) + b &= f'_1 \\2a_2(x_2 - z) + b &= f'_2\end{aligned}$$

Možnosti polohy bodu zleponia z :

Stredový splajn: $z = \frac{x_1+x_2}{2}$

Minimový splajn: $z = \bar{x}, \phi'(z) = b = 0$