

Metody numeryczne

Sprawozdanie 10

Minimalizacja wartości funkcji metodą interpolacji kwadratowej Powella

Mateusz Górczany

Maj 2020

1 Wstęp teoretyczny

Niech $f(x)$ będzie funkcją, dla której przeprowadzana jest minimalizacja, a x^* argumentem, dla którego funkcja przyjmuje minimum lokalne.

Minimalizacja wartości funkcji metodą interpolacji kwadratowej *Powella* polega na znalezieniu funkcji kwadratowej $p_2(x)$, będącej przybliżeniem $f(x)$, przechodzącej przez punkty:

$$\left\{ (x_1, f(x_1)), (x_2, f(x_2)), (x_3, f(x_3)) \right\}$$

Założenie: $f(x_1), f(x_2), f(x_3)$ spełniają warunek:

$$f(x_1) > f(x_2) > f(x_3).$$

Funkcja $p_2(x)$ jest zdefiniowana w następującej postaci:

$$p_2(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2), \quad (1)$$

zaś parametry a_0, a_1, a_2 określają równania:

$$\begin{cases} a_0 = f(x_1) \\ a_1 = F[x_1, x_2] = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \\ a_2 = F[x_1, x_2, x_3] = \frac{1}{x_3 - x_1} \left(\frac{f(x_3) - f(x_2)}{x_3 - x_2} - \frac{f(x_2) - f(x_1)}{x_2 - x_1} \right) = \frac{F[x_2, x_3] - F[x_1, x_2]}{x_3 - x_1}. \end{cases}$$

Na ich podstawie wzór (1) przyjmuje ostateczną postać:

$$p_2(x) = f(x_1) + F[x_1, x_2](x - x_1) + F[x_1, x_2, x_3](x - x_1)(x - x_2). \quad (2)$$

Obliczając miejsce zerowe pochodnej $p_2(x)$:

$$\begin{aligned} p_2'(x) &= \frac{dp_2(x)}{dx} = a_1 + a_2(x - x_2 + x - x_1) \\ &= F[x_1, x_2] + 2xF[x_1, x_2, x_3] - F[x_1, x_2, x_3](x_1 + x_2) \\ &= 0, \end{aligned}$$

otrzymuje się poniższe równanie określające x_m (będące przybliżeniem argumentu minimum funkcji $f(x)$):

$$\begin{aligned} x_m &= \frac{a_2(x_2 + x_1) - a_1}{2a_2} \\ &= \frac{x_2 + x_1}{2} - \frac{F[x_1, x_2]}{2F[x_1, x_2, x_3]} \\ &\approx x^*. \end{aligned}$$

(*) x_m jest minimum lokalnym jeśli $F[x_1, x_2, x_3] > 0$.

2 Opis rozwiązywanego zadania na laboratoriach

Interpolację przeprowadzono dla następujących funkcji:

$$\begin{aligned}f_1(x) &= \ln(x^5 + 3x^2 + x + 9), \quad x \in [-1.5, 1], \\f_2(x) &= x^6.\end{aligned}\tag{3}$$

Za x_1 przyjmowano następujące wartości:

$$\begin{aligned}x_1 &= \{-0.5, -0.9\} \text{ dla } f_1(x), \\x_1 &= \{1.5\} \text{ dla } f_2(x).\end{aligned}$$

Argumenty x_2, x_3 wyliczone zostały według formuły:

$$\begin{aligned}x_2 &= x_1 + h, \\x_3 &= x_1 + 2h,\end{aligned}$$

$$h = 0.01.$$

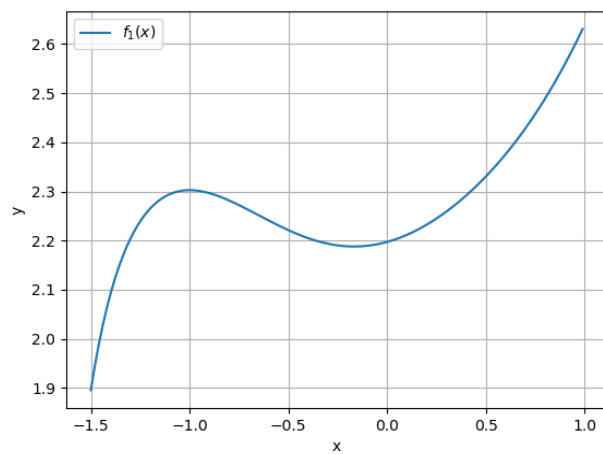
Na początku wyznaczono wartości x_1, x_2, x_3 oraz $f(x_1), f(x_2), f(x_3)$, a następnie przystąpiono do iteracyjnego obliczania x_m , za każdym razem sprawdzając warunek:

$$|x_m - x_n| < \varepsilon.$$

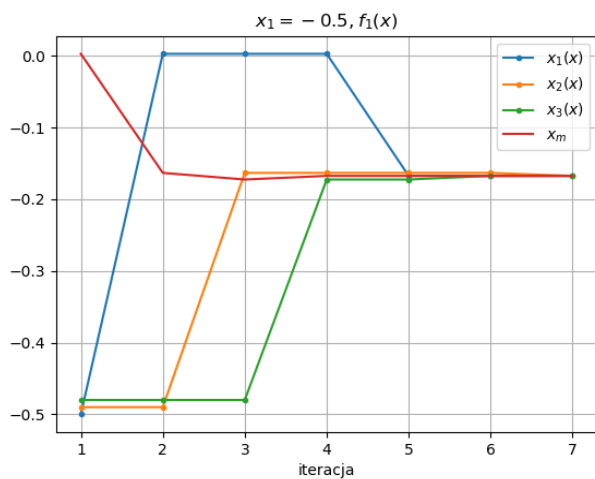
Jeśli okazał się prawdziwy to przerywano program i uznawano x_m za ostateczne przybliżenie x^* . W innym wypadku najodlegleszy punkt $x_n = \{x_1, x_2, x_3\}$ od punktu x_m przyjmował jego wartość i proces powtarzano do uzyskania satysfakcjonującego wyniku lub maksymalnej liczby iteracji. Za ε przyjęto 10^{-8} .

3 Wyniki

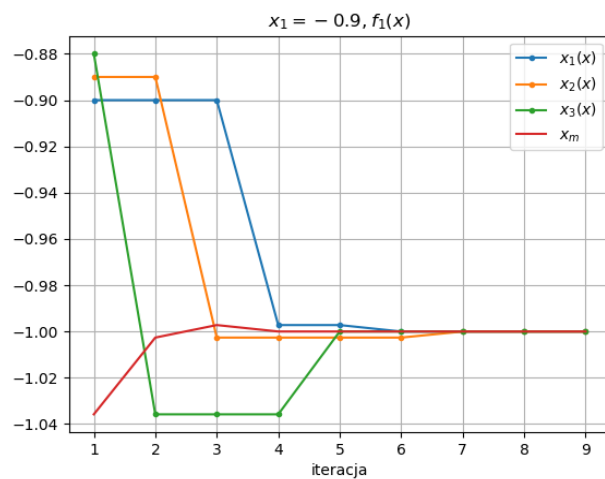
$$f_1(x) = \ln(x^5 + 3x^2 + x + 9) \downarrow$$



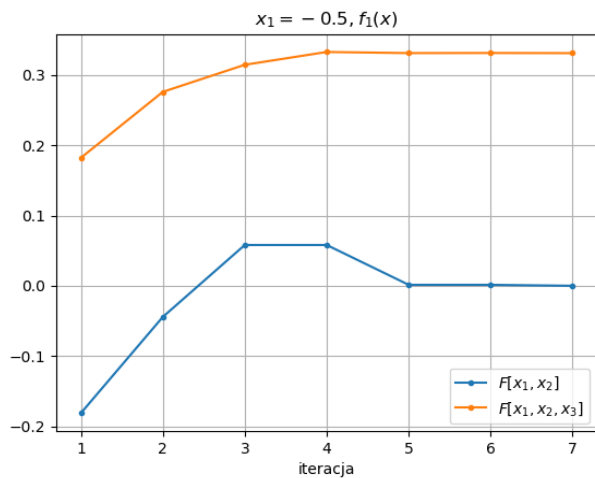
(a)



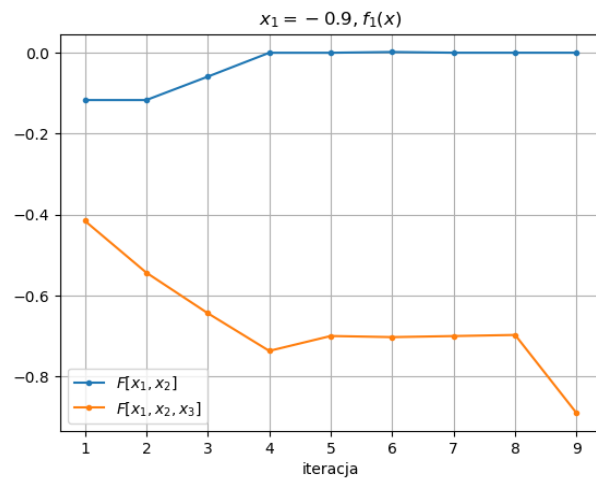
(b)



(c)

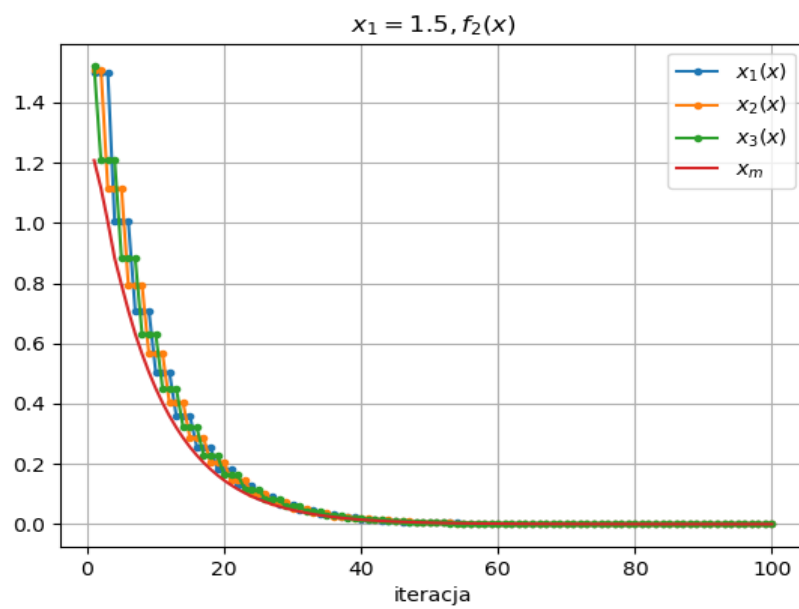


(d)

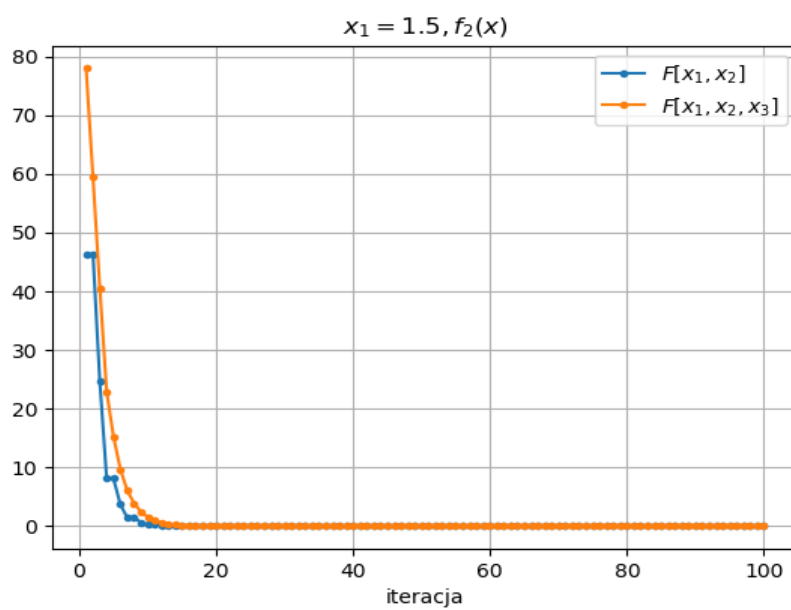


(e)

$$f_2(x) = x^6 \downarrow$$



(g)



(h)

	x_1	x_2	x_3	x_m	$F[x_1, x_2]$	$F[x_1, x_2, x_3]$
1	-0.500	-0.490	-0.480	0.003	-0.181	0.182
2	0.003	-0.490	-0.480	-0.163	-0.044	0.276
3	0.003	-0.163	-0.480	-0.172	0.058	0.315
...
7	-0.167	-0.167	-0.167	-0.167	-0.000	0.331

Table 1: $f_1(x), x_1 = -0.5$

	x_1	x_2	x_3	x_m	$F[x_1, x_2]$	$F[x_1, x_2, x_3]$
1	-0.900	-0.890	-0.880	-1.036	-0.117	-0.415
2	-0.900	-0.890	-1.036	-1.003	-0.117	-0.543
3	-0.900	-1.003	-1.036	-0.997	-0.059	-0.644
...
9	-1.000	-1.000	-1.000	-1.000	-0.000	-0.891

Table 2: $f_1(x), x_1 = -0.9$

	x_1	x_2	x_3	x_m	$F[x_1, x_2]$	$F[x_1, x_2, x_3]$
1	1.500	1.510	1.520	1.208	46.329	77.986
2	1.500	1.510	1.208	1.116	46.329	59.474
3	1.500	1.116	1.208	1.004	24.614	40.578
4	1.004	1.116	1.208	0.883	8.102	22.906
...
100	0.000019	0.000021	0.000024	0.000017	0.000	0.000

Table 3: $f_2(x), x_1 = 1.5$

4 Wnioski

Wartość $\varepsilon = 10^{-10}$ podana w treści zadania, okazała się niewystarczająca do zatrzymania pętli przed jej maksymalną iteracją. Z tego powodu przyjęto za nią 10^{-8} , co prezentują powyższe wykresy. Jeśliby zwiększyć ε , to punkt x_m w kolejnych iteracjach przyjmuje wartość "nan", gdyż iloczyn $\frac{F[x_1, x_2]}{2F[x_1, x_2, x_3]} = \frac{0}{2*0}$.

Wartość, do której zbiegają krzywe na wykresie (b) to min. lokalne, zaś na wykresie (c) to maksimum lok. Dzieje się tak, gdyż iloczyn $F[x_1, x_2, x_3]$, w drugim przypadku, jest < 0 , a więc warunek (*) nie jest spełniony.

Metoda *Powella* jest wolnozbieżna dla funkcji $f_2(x)$, gdyż iloczyny $F[x_1, x_2]$ i $F[x_1, x_2, x_3]$ przyjmują wartości bliskie zeru, stąd kolejne punkty x_m są oddalone od siebie o mały krok. W tym przypadku wartość drugiej pochodnej nie decyduje o istnieniu minimum, gdyż jej przybliżenie ($F[x_1, x_2, x_3]$) jest większe od zera i zbiega do niego w kolejnych iteracjach, a mimo to punkt $x^* = 0$ nim jest. Aby zmniejszyć liczbę iteracji dla tej funkcji, należy zwiększyć wartość h do ok. 0.75 i zmienić jego znak na (-), bądź zwiększyć ε , np. do 10^{-5} .