

Numeryczne rozwiązywanie równań nieliniowych

Wstęp

Wiele zjawisk fizycznych opisują zależności nieliniowe. Niektóre z nich można linearyzować z zachowaniem wystarczającej dokładności i analizować metodami z poprzednich rozdziałów. Gdy linearyzacja problemu nie jest możliwa lub prowadzi do zbyt dużego błędu, należy zastosować metodę przeznaczoną do równań nieliniowych.

Rozwiązania równań nieliniowych wyznacza się zwykle metodami iteracyjnymi. Obliczenia kończy się po spełnieniu kryterium zbieżności z tolerancją ε . W rozdziale omówiono następujące metody [1, 2]:

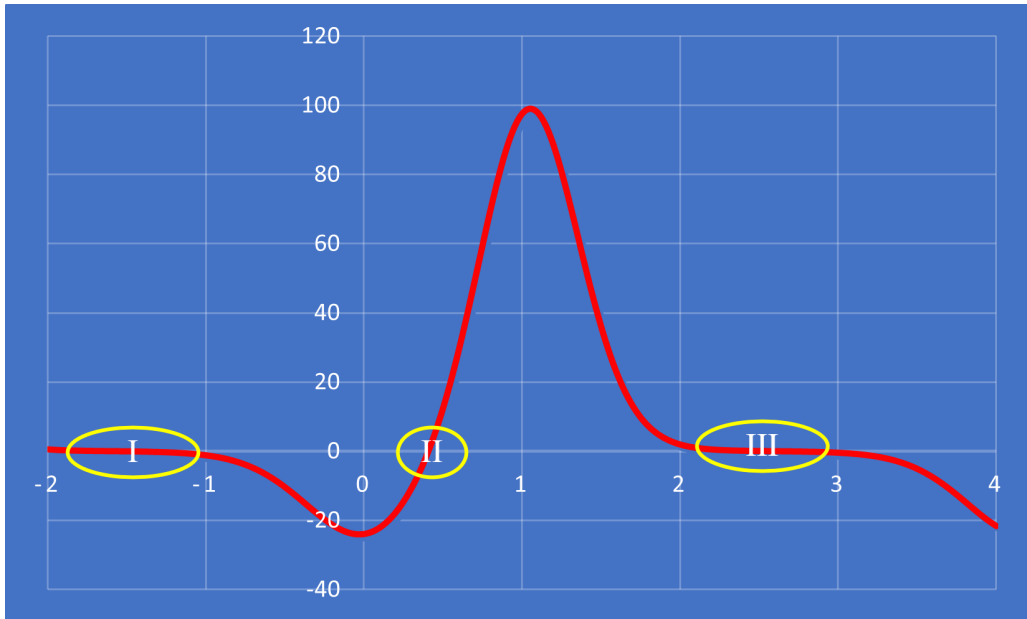
1. bisekcji
2. reguła falsi
3. siecznych
4. stycznych
5. iteracji prostej

Numeryczne rozwiązywanie równania nieliniowego powyższymi metodami iteracyjnymi realizowane jest zazwyczaj w ściśle określonym przedziale (a, b) , w którym oczekiwane jest występowanie poszukiwanego rozwiązania.

Ważnym kryterium determinującym szybkość i dokładność odnalezienia rozwiązania równania jest dopuszczalny błąd obliczeń ε równania $f(x) = 0$. W najprostszym rozwiązaniu kryterium granicznego poszukiwania rozwiązania sprawdza się warunek $f(x_k) \leq \varepsilon$. Jeżeli warunek jest spełniony uznaje się wartość x_k za rozwiązanie, jeżeli nie, kontynuuje się obliczenia wyznaczając kolejne przybliżenie. Należy jednak zwrócić uwagę na rozpiętość poprawnego rozwiązania dla wybranego błędu dopuszczalnego. Zagadnienie to najlepiej rozpatrywać na konkretnym przypadku równania nieliniowego (1).

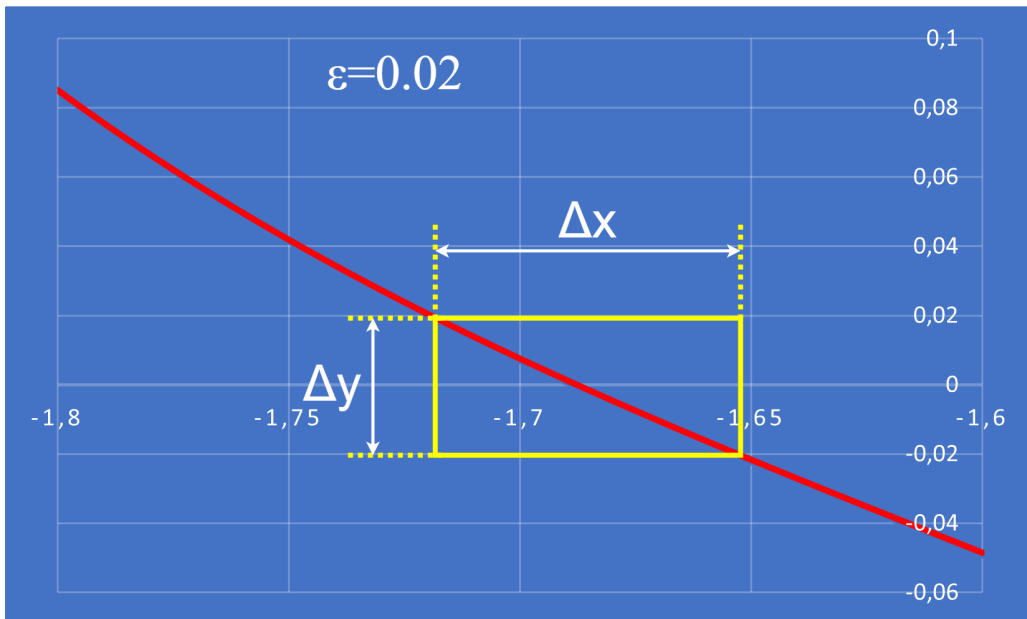
$$f(x) = 10^{2 \cdot \sin(1.5 \cdot x)} - 5^{2 \cdot \cos(1.5 \cdot x)} \quad (1)$$

Charakterystykę pokazującą przebieg funkcji w przedziale $(-2, 4)$, jak pokazano to na ryc. 1.



Ryc. 1. Przebieg przykładowej funkcji nieliniowej

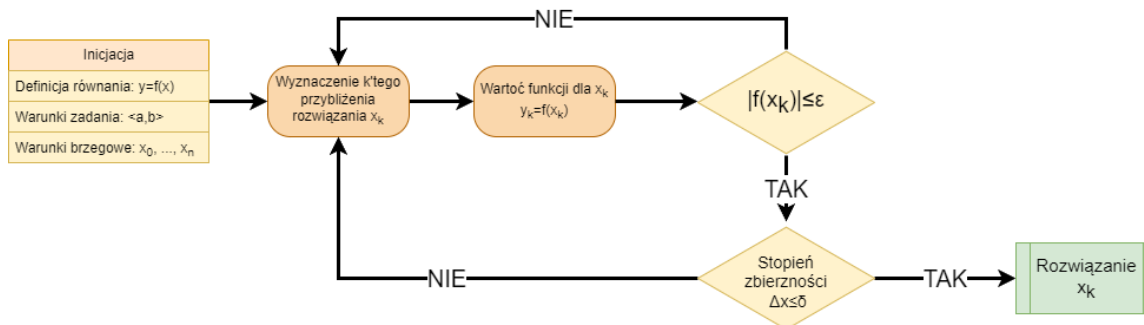
W analizowanym przedziale $(-2, 4)$ funkcja ma trzy miejsca zerowe, oznaczone na ryc. 1 jako obszary *I*, *II* i *III*. W obszarach *I* i *III* małe nachylenie funkcji powoduje, że warunek $|f(x)| \leq 0.02$ dopuszcza szeroki zakres wartości x , z błędem sięgającym około $\varepsilon_x = \pm 0.04$.



Ryc. 2. Otoczenie *I* (i *III*) miejsca zerowego

Dlatego kryterium oparte wyłącznie na wartości funkcji należy uzupełnić warunkiem dotyczącym zmiany argumentu. Można jednocześnie wymagać $|f(x_k)| \leq \varepsilon_f$ oraz $|x_k - x_{k-1}| \leq \varepsilon_x$. Taki zapis rozdziela tolerancję residuum równania od tolerancji samego rozwiązania.

Wszystkie omawiane w ćwiczeniu metody rozwiązywania równań nieliniowych opierają się na tym samym algorytmie iteracyjnym, pokazany na ryc. 3.



Ryc. 3. Uogólniony algorytm rozwiązywania równania nieliniowego metodą krokową.

1 Metoda bisekcji – połowienia

Podstawową metodą iteracyjną rozwiązywania równań nieliniowych jest metoda **bisekcji**. Przy spełnionych warunkach metody:

- w przedziale $\langle a, b \rangle$ jest jedno miejsce zerowe,
- na krańcach przedziału funkcja przyjmuje wartości o przeciwnych znakach $f(a) \cdot f(b) < 0$,

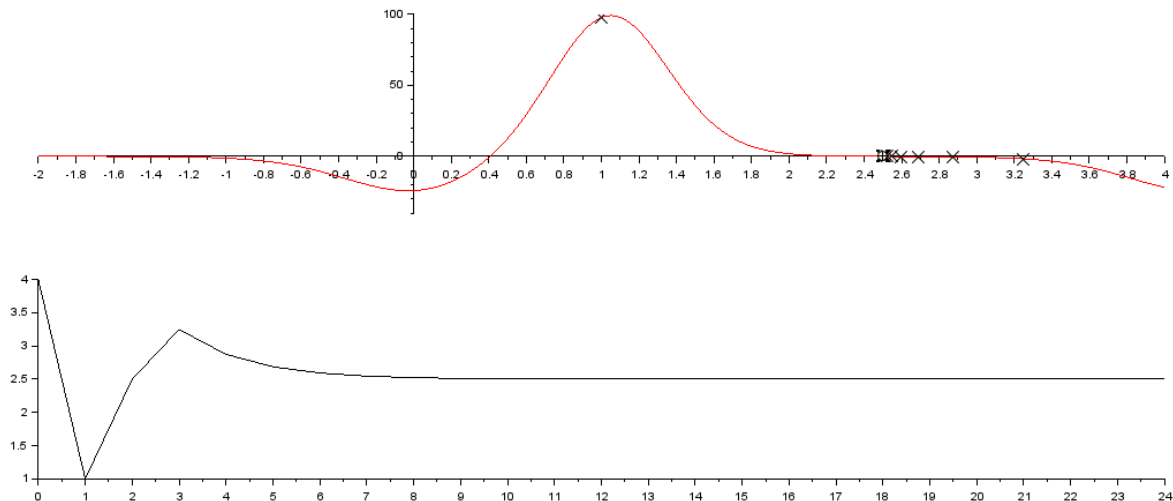
metoda gwarantuje zbieżność. Ze względu na metodykę dochodzenia do rozwiązania jest ona najwolniej zbieżna. Jeżeli w badanym przedziale $\langle a, b \rangle$ jest więcej jak jedno miejsce zerowe metoda może wskazać tylko jedno z nich, a w ekstremalnym przypadku, w zależności od zaprogramowanych kryteriów zbieżności, może nie wskazać żadnego lub wskaże fałszywe rozwiązanie.

Algorytm metody **bisekcji** zakłada cyklicznie realizowanie:

- podziału przedziału poszukiwań na dwie równe części $\langle a, c \rangle$ i $\langle c, b \rangle$,
- wyznaczenie wartości funkcji w punkcie połowienia i sprawdzenie czy $|f(c)| \leq \varepsilon$,
- po spełnieniu warunku punkt c staje się przybliżonym rozwiązaniem równania,
- przy warunku niespełnionym należy określić w którym z dwóch nowych przedziałów $\langle a, c \rangle$ lub $\langle c, b \rangle$ znajduje się rozwiązanie równania nieliniowego poprzez wykonanie testów $f(a) \cdot f(c) < 0$ i $f(c) \cdot f(b) < 0$,
- przedział, w którym zachodzi zmiana znaku, staje się nowym przedziałem poszukiwania i cykl jest powtarzany.

Rozbudowując powyższy algorytm o dodatkowe, warunkowo wykonywane testy połowienia można zaprogramować algorytm wyszukujący więcej jak jedno rozwiązanie równania nieliniowego w badanym przedziale $\langle a, b \rangle$.

Na podstawie powyższych kroków opracowano funkcję `bisekcja()`. Dla funkcji (1) w przedziale $\langle -2, 4 \rangle$ odnaleziono miejsce zerowe $x \approx 2.5$ z tolerancją $\varepsilon = 10^{-7}$ po 25 iteracjach. Zbieżność pokazano na ryc. 4.



Ryc. 4. Przebieg badanej funkcji z zaznaczonym odnalezionym rozwiązaniem oraz wykres zbieżności metody

Na ryc. 4 krzyżykami zaznaczono kolejne przybliżenia rozwiązania, a kwadratem – punkt spełniający kryterium zakończenia.

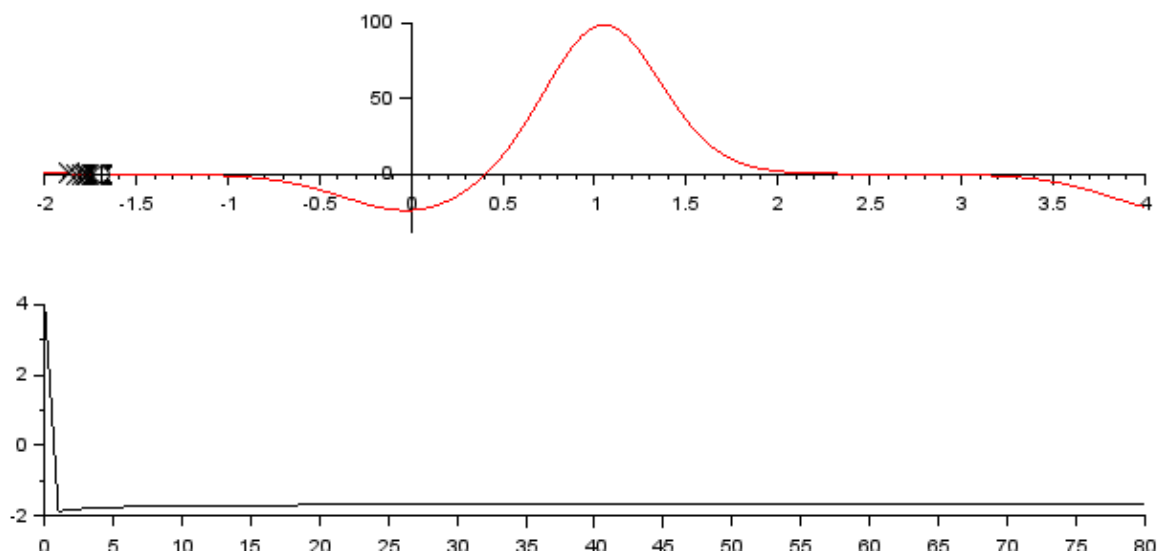
2 Metoda fałszywej pozycji (reguła falsi)

Metoda **reguła falsi** (fałszywej pozycji) jest metodą przedziałową, w której kolejne przybliżenie stanowi punkt przecięcia osi Ox z sieczną przechodzącą przez $(a, f(a))$ i $(b, f(b))$. Podobnie jak bisekcja wymaga początkowo $f(a)f(b) < 0$, dzięki czemu zachowuje przedział zawierający pierwiastek. Jeżeli $f(a) = f(b)$, wzór na punkt przecięcia ma zerowy mianownik i nie może zostać zastosowany.

Algebraicznie poszukiwane rozwiązanie można wyznaczyć z poniższej zależności.

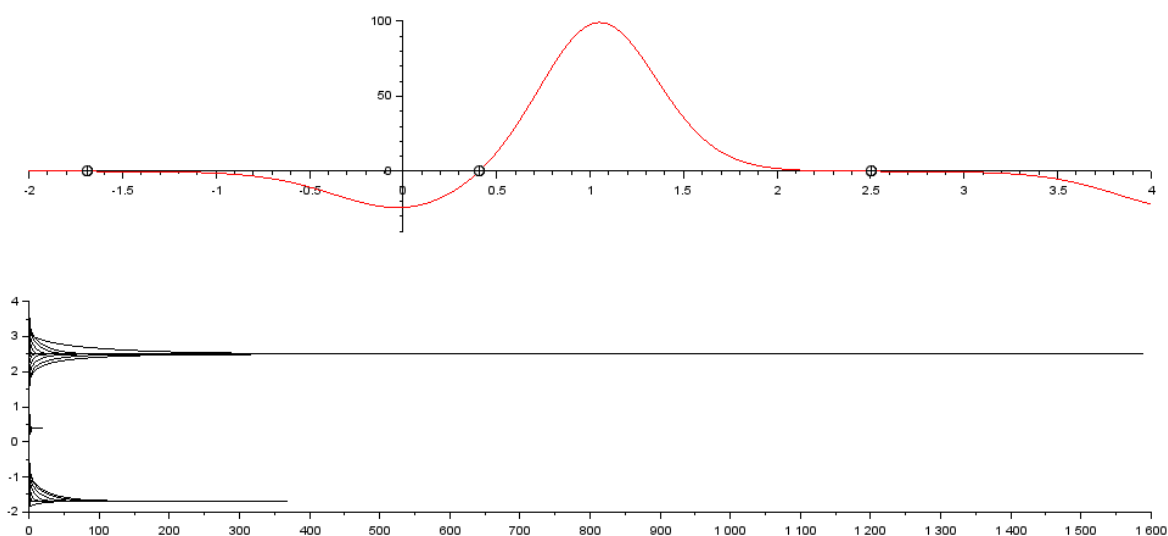
$$x_k = \frac{a \cdot f(b) - b \cdot f(a)}{f(b) - f(a)} \quad (2)$$

Algorytm różni się od bisekcji sposobem wyznaczania nowego punktu. Jeżeli przybliżenie nie spełnia kryterium, zastępuje ten koniec przedziału, dla którego wartość funkcji ma ten sam znak. Cykl powtarza się do osiągnięcia tolerancji. Dla równania (1) otrzymano $x \approx -1.69$ po 81 iteracjach. Odnalezienie pozostałych miejsc zerowych wymaga zmiany przedziału albo wcześniejszego podziału całego zakresu na podprzedziały ze zmianą znaku.



Ryc. 5. Przebieg badanej funkcji z zaznaczonym odnalezionym rozwiązaniem oraz wykres zbieżności metody

Skuteczność metody zależy od doboru przedziału początkowego i przebiegu funkcji. Klasyczna metoda fałszywej pozycji zachowuje przedział izolacji pierwiastka, jeżeli na jego końcach funkcja ma przeciwne znaki. Aby znaleźć wiele miejsc zerowych, można podzielić badany zakres na mniejsze przedziały i zastosować metodę osobno tam, gdzie wykryto zmianę znaku.



Ryc. 6. Wyszukiwanie wszystkich miejsc zerowych

Statystykę wyszukiwania zestawiono w tabeli pokazanej poniżej.

Tab. 1. Statystyka wyszukiwania rozwiązania funkcji nieliniowej w przedziale $\langle -2, 4 \rangle$ w skoku $\Delta x = 0.2$ metodą **regula fałsi**

Lp.	x_k	n_{mini} i przedział	n_{max} i przedział
1	-1.69	6 dla $\langle -1.6, -1.4 \rangle$	370 dla $\langle -0.4, -0.2 \rangle$
2	0.41	6 dla $\langle 0.2, 0.4 \rangle$	21 dla $\langle 0, 0.2 \rangle$
3	2.5	6 dla $\langle 2.6, 2.8 \rangle$	1591 dla $\langle 1, 1.2 \rangle$

3 Metoda siecznych

Metoda siecznych wykorzystuje prostą przechodzącą przez dwa ostatnie przybliżenia. Podawany przedział $\langle a, b \rangle$ dostarcza dwóch pierwszych punktów, ale później metoda nie zachowuje przedziału ani nie wymaga zmiany znaku. Z tego powodu może znaleźć rozwiązanie poza zakresem początkowym lub stać się rozbieżna.

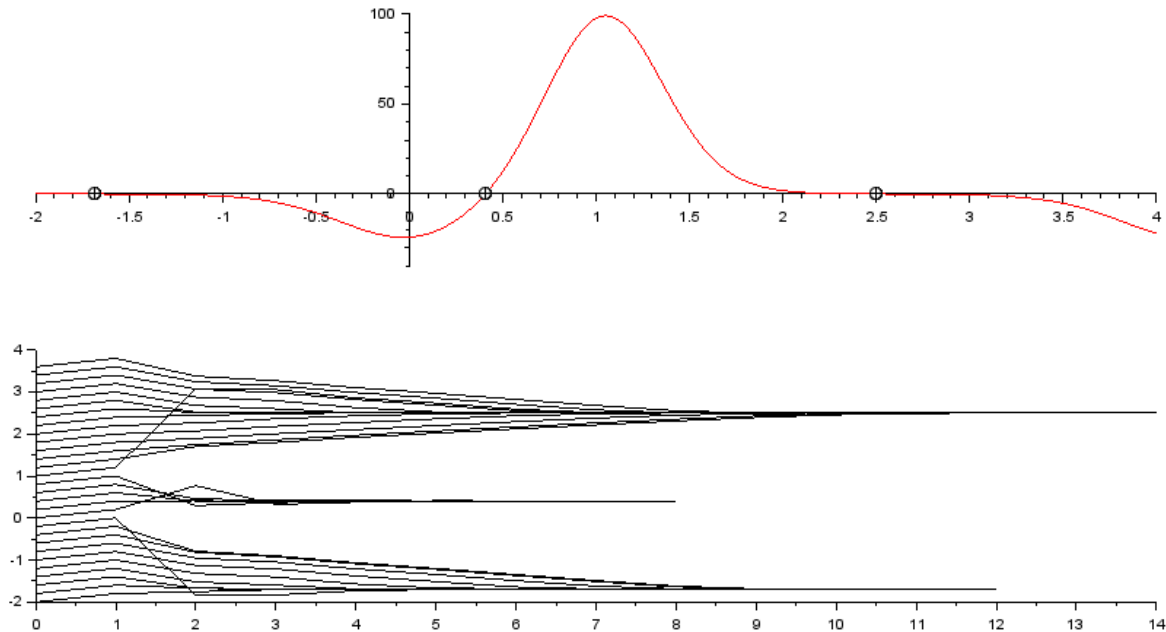
Algorytm metody siecznych składa się z trzyetapowego cyklu powtarzanego do spełnienia kryterium zbieżności:

- wyznaczenie kolejnego przybliżenia rozwiązania x_{k+1} na podstawie wcześniejszych dwóch przybliżeń x_{k-1} i x_k
- dodanie wyznaczonego punktu do listy przybliżonych rozwiązań funkcji nieliniowej
- sprawdzenie warunku zbieżności $|f(x_{k+1})| \leq \varepsilon$; dla warunku niespełnionego cykl jest powtarzany, a dla spełnionego x_{k+1} jest przybliżeniem rozwiązania równania nieliniowego

Kolejne przybliżenie można wyznaczyć z poniższej zależności albo z równania (2), odpowiednio podstawiając dwa ostatnie punkty.

$$x_{k+1} = x_k - \frac{f(x_k) \cdot (x_k - x_{k-1})}{f(x_k) - f(x_{k-1})} \quad (3)$$

Na podstawie algorytmu i zależności (3) opracowano skrypt wyznaczający miejsca zerowe funkcji (1) w przedziale $\langle -2, 4 \rangle$. Wyniki pokazano na ryc. 7, a statystyki zestawiono w Tab. 2.



Ryc. 7. Wyszukiwanie wszystkich miejsc zerowych metodą siecznych

Tab. 2. Statystyka wyszukiwania rozwiązań funkcji nieliniowej w przedziale $\langle -2, 4 \rangle$ ze skokiem $\Delta x = 0.2$ metodą siecznych.

Lp.	x_k	n_{mini} i przedział	n_{max} i przedział
1	-1.69	4 dla $\langle -1.8, -1.6 \rangle$	11 dla $\langle -0.6, -0.4 \rangle$
2	0.41	4 dla $\langle 0.2, 0.4 \rangle$	7 dla $\langle 0, 0.2 \rangle$
3	2.5	4 dla $\langle 2.4, 2.6 \rangle$	13 dla $\langle 1.2, 1.4 \rangle$

Można zauważyć, że niezależnie od wyboru punktu startowego metoda jest zdecydowanie szybsza od omówionych wcześniej metod **bisekcji** i **regula falsi**. Po ustawieniu punktów granicznych poszukiwania rozwiązania $\langle -2, 4 \rangle$ jako pierwszych przybliżeń odnajdywane jest pierwsze miejsce zerowe -1.69 po 7 krokach.

4 Metoda stycznych

Metoda **stycznych**, nazywana także metodą Newtona-Raphsona, ma lokalnie zbieżność kwadratową, jeżeli pierwiastek jest prosty, funkcja jest dostatecznie gładka, a punkt początkowy leży odpowiednio blisko rozwiązania. Dla wyboru punktu startowego w przedziale często stosuje się następujące warunki wystarczające:

- w przedziale poszukiwania musi posiadać dokładnie jedno rozwiązanie,
- ma różne znaki na krańcach przedziału $f(a)f(b) < 0$,
- pierwsza i druga pochodna w zadanym przedziale mają stały znak.

W punkcie x_k funkcję zastępuje się jej liniowym przybliżeniem, czyli styczną do wykresu. Punkt przecięcia tej stycznej z osią Ox wyznacza kolejne przybliżenie x_{k+1} . Metoda wymaga jednego punktu początkowego i nie gwarantuje pozostania w zadanym przedziale.

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad (4)$$

Zależność (4) wymaga obliczenia pierwszej pochodnej w punkcie x_k . Dla funkcji testowej (1) pochodna ma postać:

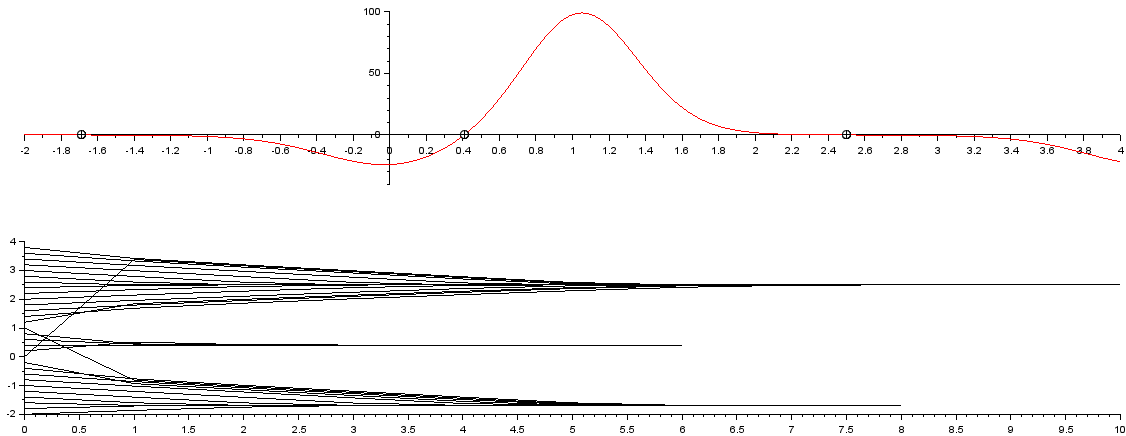
$$f'(x) = 3 \left[\log(10) \cos(1,5x) 10^{2 \sin(1,5x)} + \log(5) \sin(1,5x) 5^{2 \cos(1,5x)} \right] \quad (5)$$

W środowisku **Scilab** pochodną funkcji w punkcie można przybliżyć funkcją `numderivative()`:

```
dy=numderivative(funkcja,xk);
```

Argument `funkcja` jest uchwytym do funkcji nieliniowej, a `xk` – punktem, w którym oblicza się pochodną. Funkcja `diff()` służy natomiast do różnicowania tablic próbek; iloraz `diff(y)./diff(x)` daje wartości na przedziałach między kolejnymi punktami.

Dla pochodnej analitycznej i numerycznej uzyskano zgodne wyniki. Skanując przedział $\langle -2, 4 \rangle$ z krokiem $\Delta x = 0.2$, odnaleziono wszystkie miejsca zerowe po 4–10 iteracjach, jak pokazano na ryc. 8. Można zauważyć, że im mniejsze jest nachylenie funkcji w pobliżu miejsca zerowego, tym bardziej wybór punktu początkowego wpływa na przebieg i zbieżność iteracji.

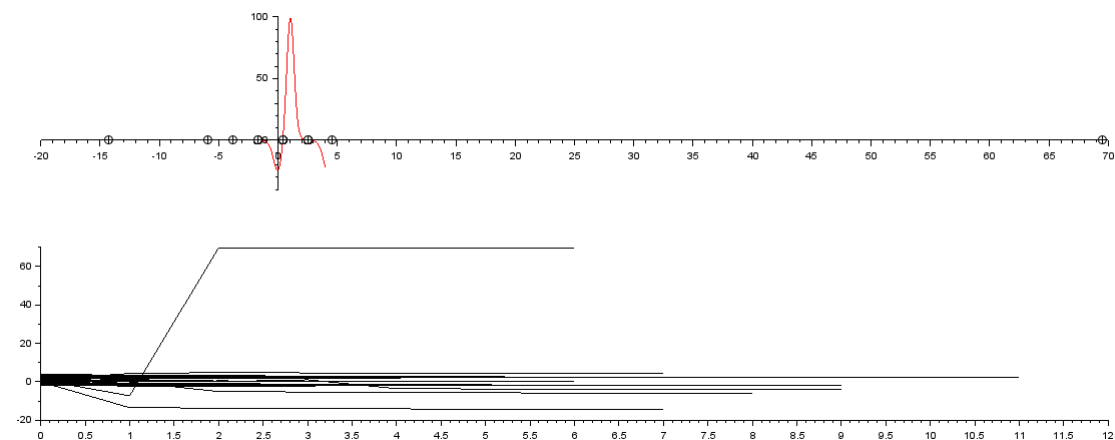


Ryc. 8. Wyszukiwanie wszystkich miejsc zerowych metodą stycznych

W zależności od punktu początkowego metoda może odnaleźć:

- najbliższe miejsce zerowe,
- inne miejsce zerowe z analizowanego przedziału,
- miejsce zerowe spoza przedziału,
- nie odnaleźć żadnego miejsca zerowego,

, jak pokazano na ryc. 9.



Ryc. 9. Wyszukiwanie wszystkich miejsc zerowych metodą stycznych, także poza przedziałem poszukiwania

5 Metoda iteracji prostej

W metodzie **iteracji prostej** przybliżoną wartość rozwiązania równania nieliniowego $f(x) = 0$ wyznacza się z zależności $\widehat{x}_k = g(x_{k-1})$ wyznaczonej poprzez przekształcenie równania właściwego. Realizowane jest to na analogicznych zasadach jak przy metodzie iteracyjnej ogólnej stosowanej dla układów równań liniowych.

Warunkiem podstawowym jest występowanie w równaniu nieliniowym kilku (przynajmniej dwóch) członów z niewiadomą. Możliwe jest więc poprzez wyłączenie niewiadomej z każdego z tych członów uzyskać odpowiadającą mu postać równania iteracyjnego. Aby równanie iteracyjne mogło być zbieżne do rozwiązania równania nieliniowego musi spełniać poniższe warunki:

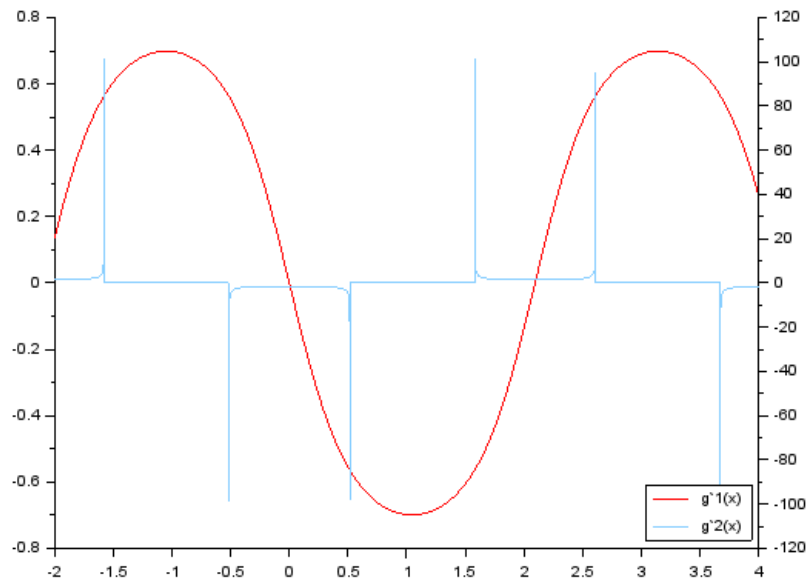
- $|g'(x)| < 1$
- $|g(x_{k+1}) - g(x_k)| < |g(x_k) - g(x_{k-1})|$

Dla równania nieliniowego (1) można zapisać dwa równania iteracyjne:

$$g_1(x) = \frac{2}{3} \cdot \sin\left(\frac{1}{2} \log_{10} 5^{2 \cdot \cos(\frac{3}{2} \cdot x)}\right) \quad (6)$$

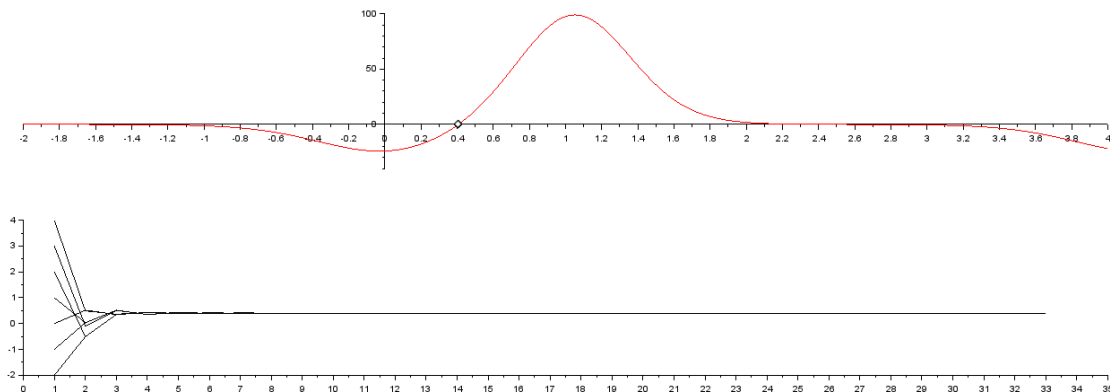
$$g_2(x) = \frac{2}{3} \cdot \cos\left(\frac{1}{2} \log_5 10^{2 \cdot \sin(\frac{3}{2} \cdot x)}\right) \quad (7)$$

Dla obu przekształceń sprawdzono warunek zbieżności, pokazany na ryc. 10. Jest on spełniony tylko dla $g_1(x)$.



Ryc. 10. Przebieg zmian pierwszych pochodnych równań iteracyjnych $g_1(x)$ i $g_2(x)$

Przeprowadzona dla równania iteracyjnego (6) symulacja niezależnie od wartości pierwszego przybliżenia x_k zwraca to samo rozwiązanie o wartości 0.41 po około 31 krokach (liczba iteracji niemal identyczna dla każdego punktu startowego x_0), jak pokazano na ryc. 11.



Ryc. 11. Wyszukiwanie wszystkich miejsc zerowych metodą iteracji prostej dla równania iteracyjnego $g_1(x)$

Skuteczność iteracji prostej zależy od wyboru funkcji $g(x)$ i punktu początkowego. W środowisku **Scilab** ogólne równania i układy nieliniowe rozwiązuje funkcja `fsolve()`, której podstawowe argumenty to punkt startowy x_0 oraz funkcja definiująca residuum. Dla równania (1) i dwóch punktów startowych funkcja zwróciła dwa różne miejsca zerowe: $x_1 \approx -1.69$ i $x_2 \approx 2.5$.

Zadania

Rozpatrzyc obwód złożony ze źródła napięcia U_s , rezystora R i diody opisanej równaniem Shoc-kleya. Każdy student przyjmuje parametry z wiersza odpowiadającego numerowi zestawu. Punkt pracy spełnia równanie

$$f(U_D) = I_s \left[\exp \left(\frac{U_D}{nU_T} \right) - 1 \right] - \frac{U_s - U_D}{R} = 0.$$

Zestaw	U_s [V]	R [Ω]	I_s [pA]	n
1	5,0	1000	1,0	1,50
2	3,3	680	1,0	1,50
3	12,0	2200	1,0	1,50
4	5,0	470	1,0	1,50
5	9,0	1500	2,0	1,60
6	6,0	820	0,5	1,40
7	15,0	3300	1,5	1,70
8	4,5	560	5,0	1,80
9	10,0	1000	0,8	1,30
10	7,5	1800	3,0	1,55
11	5,5	1200	1,2	1,45
12	18,0	3900	0,7	1,65
13	2,5	390	4,0	1,75
14	8,0	680	1,0	1,35
15	11,0	2700	2,5	1,60

Zestaw	U_T [V]	x_0 [V]	x_1 [V]
1	0,02585	0,60	0,90
2	0,02585	0,58	0,88
3	0,02585	0,62	0,92
4	0,02585	0,62	0,92
5	0,02585	0,60	0,95
6	0,02585	0,58	0,90
7	0,02585	0,62	0,98
8	0,02585	0,55	0,90
9	0,02585	0,58	0,88
10	0,02585	0,58	0,94
11	0,02585	0,58	0,90
12	0,02585	0,62	1,00
13	0,02585	0,52	0,85
14	0,02585	0,58	0,90
15	0,02585	0,60	0,96

We wszystkich zestawach przyjąć przedział metod przedziałowych $[0, 1]$ V, tolerancję 10^{-10} i limit 100 iteracji.

1. Narysować charakterystykę prądowo-napięciową diody oraz prostą obciążenia. Na podstawie wykresu wskazać przedział zawierający punkt pracy.
2. Wyznaczyć punkt pracy metodami bisekcji i fałszywej pozycji. Zastosować kryterium oparte na szerokości przedziału i wartości $|f(U_D)|$.

3. Wyznaczyć rozwiązanie metodami siecznych i Newtona. Przeanalizować wpływ punktów startowych na zbieżność. Wyniki zweryfikować funkcją `fsolve()`.
4. Porównać liczbę iteracji i końcowe residua. Dla swojego zestawu powtórzyć obliczenia dla $0,8U_s$, U_s i $1,2U_s$ oraz przedstawić zmianę punktu pracy.

Zadanie rozszerzające. Zbadać wpływ temperatury, uwzględniając zmianę napięcia termicznego U_T i prądu nasycenia I_s , albo przeanalizować obwód zawierający drugi element nieliniowy.

Bibliografia

- [1] B. Mochracki E. Majchrzak. *Metody numeryczne. Podstawy teoretyczne, aspekty praktyczne i algorytmy*. Wydawnictwo Politechniki Śląskiej, 2004.
- [2] J. Povstenko. *Wprowadzenie do metod numerycznych*. Warszawa: Akademicka oficyna wydawnicza EXIT, 2002.