

Rozwiązywanie równań nieliniowych

Wykład 1 i 2

- Metoda bisekcji, Newtona i siecznych
- Pierwiastki wielomianów
- Wzór Taylora i jego warianty
- Układy równań nieliniowych

Wprowadzenie.

Podstawowym zagadnieniem znanym z algebry jest znajdowanie pierwiastków równania

$$f(x) = 0,$$

tzn. takiego argumentu, który daje wartość zero (spełnia równanie). Mówiąc bardziej dokładnie, jeżeli funkcja jest zdefiniowana jako

$$y = f(x),$$

szukamy takiego argumentu α , że

$$f(\alpha) = 0.$$

Argument α będziemy nazywać *miejscem zerowym* funkcji f lub *pierwiastkiem (rozwiązaniem)* równania $f(x) = 0$.

Twierdzenie 1. (Własność Darboux funkcji ciągłych)
Funkcja ciągła f w przedziale $[a, b]$ przyjmuje w nim wszystkie wartości zawarte między $f(a)$ i $f(b)$.

Metoda bisekcji (połowienia przedziału)

Niech $y = f(x)$ będzie daną funkcją. Przypuśćmy, że możemy stwierdzić następujący fakt

$$f(a)f(b) < 0.$$

Oznacza to, że f jest ujemne w jednym punkcie (a lub b) i dodatnie w drugim. Jeżeli ponadto f jest funkcją ciągłą, wówczas na podstawie własności Darboux wiemy, że istnieje taki argument pomiędzy a i b (oznaczymy go α) dla którego f przyjmuje wartość zero. Innymi słowy istnieje pierwiastek $\alpha \in [a, b]$ dla którego mamy

$$f(\alpha) = 0.$$

(Uwaga: W przedziale $[a, b]$ może być więcej niż jeden pierwiastek.)

Użyjmy teraz powyższego pomysłu do znalezienia pierwiastka α . Niech c będzie środkiem przedziału $[a, b]$, tzn.

$$c = \frac{1}{2}(a + b).$$

Przyjrzyjmy się iloczynowi $f(a)f(c)$.

Mamy trzy możliwości:

- 1.** $f(a)f(c) < 0$. Oznacza to, że ten pierwiastek (być może więcej niż jeden) znajduje się pomiędzy a i c , tzn. $\alpha \in [a, c]$;
- 2.** $f(a)f(c) = 0$. Wiemy już, że $f(a) \neq 0$, zatem musi być $f(c) = 0$. Oznacza to, że znaleźliśmy pierwiastek, mianowicie $\alpha = c$.
- 3.** $f(a)f(c) > 0$. Oznacza to, że pierwiastek musi leżeć w drugiej połowie przedziału, tzn. $\alpha \in [c, b]$.

Jeżeli zachodzi przypadek 1 lub 3 mamy wówczas pierwiastek zlokalizowany w przedziałach $[a, c]$ lub $[c, b]$, które są o połowę krótsze od wyjściowego przedziału $[a, b]$.

Jeżeli powtórzymy ten proces ponownie otrzymamy krótszy o połowę przedział lokalizujący pierwiastek.

Kontynuując to postępowanie możemy zlokalizować pierwiastek z dowolną, wybraną z góry dokładnością, tzn. otrzymać przedział lokalizujący dowolnie małej długości.

Kryterium stopu (zakończenia obliczeń)

Jeżeli zachodzi przypadek 2, to pierwiastek jest znaleziony. W praktyce ten przypadek ma rzadko miejsce. W związku z tym nie prowadzimy obliczeń, aż do otrzymania

$$f(c) = 0,$$

ale dopuszczamy pewną tolerancję, np. kończymy obliczenia, gdy mamy

$$|f(c)| < 10^{-8}.$$

Innym kryterium zakończenia obliczeń może być dostatecznie mała długość przedziału lokalizacyjnego lub osiągnięcie maksymalnej dopuszczonej liczby powtórzeń.

Ćwiczenie 1. Wyznaczyć miejsce zerowe funkcji $f(x) = 2 - e^x$ zaczynając poszukiwania od przedziału $[0, 1]$, z dokładnością do 10^{-4} .

Ćwiczenie 2. Wyznaczyć pierwiastek równania $e^x = \sin x$ najbliższy 0 z dokładnością do 10^{-4} .

Zbieżność i błąd metody bisekcji

Twierdzenie 2. Niech $[a_0, b_0] = [a, b]$ będzie początkowym przedziałem z $f(a)f(b) < 0$ i niech przedziały $[a_1, b_1], [a_2, b_2], \dots$ będą utworzone metodą bisekcji. Zdefiniujemy przybliżony pierwiastek jako

$$x_n = c_n = \frac{a_{n-1} + b_{n-1}}{2}.$$

Wtedy istnieje dokładny pierwiastek $\alpha \in [a, b]$ funkcji f taki, że

$$|\alpha - x_n| \leq \left(\frac{1}{2}\right)^n (b - a).$$

Ponadto, aby osiągnąć dokładność

$$|\alpha - x_n| \leq \varepsilon,$$

wystarczy wziąć

$$n \geq \frac{\log(b - a) - \log \varepsilon}{\log 2}.$$

Ćwiczenie 3. Niech metoda bisekcji startuje od przedziału $[50, 63]$. Ile co najwyżej kroków trzeba wykonać, aby otrzymać pierwiastek z dokładnością 10^{-12} .

Metoda Newtona (metoda stycznych)

Klasyczna metoda znajdowania pierwiastków funkcji. Historycznie po raz pierwszy użyta przez Izaaca Newtona w 1669 roku, choć sama idea była znana już wcześniej Josephowi Raphsonowi (dlatego metoda bywa nazywana metodą Newtona-Raphsona). Starożytni Babilończycy znali algorytm przybliżania pierwiastków kwadratowych z liczby dodatniej, co do istoty oparty na tym rozumowaniu.

Metoda Newtona z definicji zaczyna od przyjęcia (wytypowania) pierwszego przybliżenia x_0 pierwiastka α funkcji f i polega na rekurencyjnym stosowaniu wzoru

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

dla $n \geq 0$.

Kryterium stopu

Do pomiaru zbieżności metody możemy użyć wielkości $|x_n - x_{n-1}|$. Zazwyczaj kończymy iterowanie, kiedy ta wielkość jest odpowiednio mała, np.

$$|x_n - x_{n-1}| \leq 10^{-12}.$$

W praktyce może się zdarzyć, że mamy $|x_n - x_{n-1}|$ małe, a jednocześnie x_n jest niezbyt bliskie α . Może to mieć miejsce, gdy $f'(x_n)$ jest bardzo duże w porównaniu z $f(x_n)$. Z tego powodu często dodaje się składnik zapewniający, że sama wartość funkcji jest również mała, tzn. kończymy obliczenia kiedy

$$|f(x_n)| + |x_n - x_{n-1}| \leq 10^{-12}.$$

Innym kryterium może być osiągnięcie maksymalnej dopuszczalnej liczby iteracji.

Ćwiczenie 4. Zastosować metodę Newtona do znalezienia pierwiastka funkcji $f(x) = 2 - e^x$ z dokładnością do 10^{-8} . Przyjąć pierwsze przybliżenie $x_0 = 0$.

Ćwiczenie 5. Stosując metodę Newtona znaleźć ujemny pierwiastek funkcji $g(x) = e^x - 1.5 - \arctan x$ z dokładnością do 10^{-12} .

Zbieżność i błąd metody Newtona

Twierdzenie 3. *Niech α będzie pojedynczym miejscem zerowym funkcji f i niech jej druga pochodna f'' będzie ciągła. Wtedy istnieje takie otoczenie punktu α i taka stała C , że jeśli metoda Newtona startuje z tego otoczenia, to jej kolejne punkty są coraz bliższe α i takie, że*

$$|x_{n+1} - \alpha| \leq |C(x_n - \alpha)^2|$$

dla $n \geq 0$.

W pewnych przypadkach metoda Newtona jest zbieżna dla dowolnego punktu początkowego.

Twierdzenie 4. *Jeżeli $f \in C^2(\mathbb{R})$ jest rosnąca, wypukła i ma miejsce zerowe, to jest ono jedyne, a metoda Newtona daje ciąg do niego zbieżny dla dowolnego punktu początkowego.*

Ćwiczenie 6. Obliczyć $\sqrt{26}$. Porównać ze wzorem Herona:

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{R}{x_n} \right).$$

Metoda siecznych

Wadą metody Newtona jest to, że wymaga ona wzoru na pochodną funkcji f . W praktyce zdarzają się sytuacje w których obliczenie pochodnej może być trudne albo nieopłacalne. Jednym ze sposobów poradzenia sobie z tym problemem jest zastąpienie pochodnej we wzorze Newtona jej przybliżeniem, mianowicie

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}.$$

Ta równość przybliżona, która wynika wprost z definicji pochodnej, prowadzi do metody siecznych danej wzorem rekurencyjnym

$$x_{n+1} = x_n - f(x_n) \left[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \right]$$

dla $n \geq 1$.

Zbieżność i błąd metody siecznych

Twierdzenie 5. Niech f będzie podwójnie różniczkowalna w sposób ciągły w pewnym otoczeniu pierwiastka α , ponadto załóżmy, że $f'(x) \neq 0$ dla wszystkich x z tego otoczenia. Wtedy dla x_0 i x_1 dostatecznie bliskich α metoda siecznych daje ciąg zbieżny do α taki, że

$$\lim_{n \rightarrow \infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = 0$$

oraz

$$\lim_{n \rightarrow \infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^p} = \left(\frac{1}{2} \frac{f''(\alpha)}{f'(\alpha)} \right)^{p-1}$$

dla $p = \frac{1+\sqrt{5}}{2} \approx 1.618 \dots$

Zauważmy, że $p = \frac{1+\sqrt{5}}{2} \approx 1.62 < 2$ co oznacza, że metoda siecznych jest wolniej zbieżna od metody Newtona, ale szybciej od metody bisekcji. Zaletą metody siecznych jest to, że każdy krok tej metody wymaga obliczenia tylko jednej wartości funkcji, tzn. $f(x_n)$. W metodzie Newtona trzeba obliczyć dwie takie wartości w każdym kroku, mianowicie $f(x_n)$ i $f'(x_n)$.

W metodzie siecznych możemy zastosować te same kryteria zakończenia obliczeń jak w metodzie Newtona.

Ćwiczenie 7. Zastosować metodę siecznych do znalezienia miejsca zerowego funkcji $f(x) = 2 - e^x$ dla punktów początkowych $x_0 = 0$ i $x_1 = 1$.

Ćwiczenie 8. Znaleźć metodą siecznej najbliższy zera i dodatni pierwiastek równania $\tan x - 30x = 0$ z dokładnością do 10^{-12} .

Pierwiastki wielomianów

Wszystkie metody opisane wcześniej możemy zastosować do wielomianów. Warto jednak uwzględnić specjalną strukturę tych funkcji. Wielomianem nazywamy funkcję postaci

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0,$$

gdzie współczynniki a_k i zmienna z mogą być zespolone. Stopniem wielomianu nazywamy wykładnik najwyższej potęgi, tzn. jeżeli $a_n \neq 0$, to p ma stopień n .

Zasadniczym pytaniem przy poszukiwaniu pierwiastków wielomianów jest czy one istnieją i jeżeli istnieją, to w jakiej liczbie. Odpowiedz daje następujące twierdzenie.

Twierdzenie 6. *Wielomian stopnia n ma dokładnie n pierwiastków na płaszczyźnie zespolonej, przy czym każdy z nich liczony jest tyle razy, ile wynosi jego krotność.*

Przy poszukiwaniu pierwiastków wielomianów warto wiedzieć, jak z grubsza są one rozmieszczone na płaszczyźnie zespolonej. Odpowiedz daje twierdzenie lokalizacyjne.

Twierdzenie 7. *Wszystkie pierwiastki wielomianu stopnia n leżą w kole otwartym o środku w punkcie 0 płaszczyzny zespolonej i promieniu*

$$\rho = 1 + \frac{\max_{0 \leq k < n} |a_k|}{|a_n|}.$$

Ćwiczenie 9. *Znaleźć koło o środku w punkcie 0, zawierające wszystkie pierwiastki wielomianu*

$$p(z) = z^4 - 4z^3 + 7z^2 - 5z - 2.$$

Podobną informację o lokalizacji pierwiastków wielomianu uzyskujemy korzystając z funkcji

$$s(z) = z^n p\left(\frac{1}{z}\right) = a_n + a_{n-1}z + a_{n-2}z^2 + \cdots + a_0z^n,$$

tzn. wielomianu stopnia nie większego od n , o współczynnikach takich jak w p , ale uporządkowanych odwrotnie. Dla różnej od zera liczby zespolonej z_0 warunki $p(z_0)$ i $s\left(\frac{1}{z_0}\right)$ są równoważne. Wynika stąd następujący wniosek.

Twierdzenie 8. *Jeżeli wszystkie pierwiastki wielomianu s leżą w kole $|z| < \rho$, to wszystkie niezerowe pierwiastki wielomianu p leżą poza kołem $|z| \leq \frac{1}{\rho}$.*

Ćwiczenie 10. *Znaleźć koło o środku w punkcie 0, w którym wielomian*

$$p(z) = z^4 - 4z^3 + 7z^2 - 5z - 2$$

nie ma żadnego pierwiastka.

Wzór Taylora i jego warianty

Wzór ten dotyczy funkcji z $C^n[a, b]$. Posługujemy się nim bardzo często w analizie numerycznej.

Twierdzenie 9. *Jeśli $f \in C^n[a, b]$ i jeśli $f^{(n+1)}$ istnieje w przedziale otwartym (a, b) , to dla dowolnych punktów c i x z przedziału domkniętego $[a, b]$ mamy*

$$f(x) = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(c)(x - c)^k + E_n(x),$$

gdzie dla pewnego punktu ξ leżącego między c i x

$$E_n(x) = \frac{1}{(n + 1)!} f^{(n+1)}(\xi)(x - c)^{n+1}.$$

Wyrażenie $E_n(x)$ nazywamy *resztą Lagrange'a* wzoru Taylora. Zwrot „leżenia pomiędzy” użyty powyżej należy rozumieć tak, że albo $c < \xi < x$, albo $x < \xi < c$, w zależności od wartości c i x .

Dla $c = 0$ powyższy wzór nazywamy *wzorem Macalurina*.

Ćwiczenie 11. *Podać wzór Taylora dla*

$$f(x) = \ln x$$

przyjmując $a = 1, b = 2, c = 1$.

Ważny wariant wzoru Taylora można otrzymać zmieniając x na $x + h$ oraz c na x .

Twierdzenie 10. *Dla funkcji $f \in C^{n+1}[a, b]$ oraz dowolnych punktów x i $x + h$ z przedziału domkniętego $[a, b]$ mamy*

$$f(x + h) = \sum_{k=0}^n \frac{h^k}{k!} f^{(k)}(x) + E_n(h),$$

gdzie dla pewnego ξ leżącego między x i $x + h$

$$E_n(h) = \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(\xi).$$

Stosując powyższy wzór Taylora możemy wyprowadzić metodę Newtona. Niech f będzie funkcją, której pierwiastki chcemy wyznaczyć numerycznie. Oznaczmy przez α taki pierwiastek, a przez x jego przybliżenie. Jeśli f'' istnieje, to ze wzoru Taylora otrzymujemy

$$0 = f(\alpha) = f(x + h) = f(x) + hf'(x) + E_1(h),$$

gdzie $h = \alpha - x$. Jeżeli h jest małe (tzn. jeżeli x jest bliskie α), to resztę $E_1(h)$ możemy pominąć. Dostajemy wtedy równanie

$$f(x) + hf'(x) = 0,$$

które rozwiązujemy względem h i otrzymujemy

$$h = -\frac{f(x)}{f'(x)}.$$

Jeśli x jest przybliżeniem pierwiastka α , to jeszcze lepszym przybliżeniem tego pierwiastka jest

$$x + h = x - \frac{f(x)}{f'(x)}.$$

Startując od przybliżenia x_0 pierwiastka α w pierwszym kroku dostajemy

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Powtarzając to rozumowanie w drugim kroku otrzymujemy

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

i ogólnie

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n \geq 0).$$

Podamy teraz wariant wzoru Taylora dla funkcji dwóch zmiennych.

Twierdzenie 11. *Jeśli funkcja f należy do $C^{n+1}([a, b] \times [c, d])$ i jeśli punkty (x, y) i $(x + h, y + k)$ leżą w prostokącie $[a, b] \times [c, d] \subseteq \mathbb{R}^2$, to*

$$f(x + h, y + k) = \sum_{i=0}^n \frac{1}{i!} \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^i f(x, y) + E_n(h, k),$$

gdzie

$$E_n(h, k) = \frac{1}{(n + 1)!} \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^{n+1} f(x + \theta h, y + \theta k) \quad (0 < \theta < 1).$$

Układy równań nieliniowych

Stosując do układów równań, pomysł prowadzący do metody Newtona dla jednego równania, możemy otrzymać iteracyjną metodę numerycznego rozwiązywania układów nieliniowych. Wyprowadzenie tej metody pokażemy na przykładzie układu dwóch równań z dwiema niewiadomymi

$$\begin{cases} f_1(x_1, x_2) = 0, \\ f_2(x_1, x_2) = 0. \end{cases}$$

Pomysł polega na lineryzacji tych równań, (tzn. zastosowaniu wzoru Taylora, przy użyciu tylko liniowych członów rozwinięcia) a następnie rozwiązaniu układu w celu obliczenia poprawek.

Jeżeli (x_1, x_2) jest przybliżonym rozwiązaniem układu, to $(x_1 + h_1, x_2 + h_2)$, gdzie h_1, h_2 są obliczonymi poprawkami, powinno być jeszcze lepszym przybliżeniem. Ze wzoru Taylora dostajemy

$$\begin{cases} 0 = f_1(x_1 + h_1, x_2 + h_2) \approx f_1(x_1, x_2) + h_1 \frac{\partial f_1}{\partial x_1} + h_2 \frac{\partial f_1}{\partial x_2}, \\ 0 = f_2(x_1 + h_1, x_2 + h_2) \approx f_2(x_1, x_2) + h_1 \frac{\partial f_2}{\partial x_1} + h_2 \frac{\partial f_2}{\partial x_2}. \end{cases}$$

Pochodne cząstkowe występujące w tym układzie są obliczane w punkcie (x_1, x_2) .

Otrzymujemy stąd układ

$$\begin{cases} h_1 \frac{\partial f_1}{\partial x_1} + h_2 \frac{\partial f_1}{\partial x_2} = -f_1(x_1, x_2), \\ h_1 \frac{\partial f_2}{\partial x_1} + h_2 \frac{\partial f_2}{\partial x_2} = -f_2(x_1, x_2), \end{cases}$$

który jest liniowy względem h_1 i h_2 . Możemy zapisać go w postaci macierzowej

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = - \begin{bmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{bmatrix}.$$

Macierzą główną tego układu jest *jacobian*

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}.$$

Możemy więc zapisać

$$J \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = - \begin{bmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{bmatrix}.$$

Jeżeli macierz J jest nieosobliwa, to rozwiązanie układu istnieje i jest postaci

$$\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = -J^{-1} \begin{bmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{bmatrix}.$$

Metoda Newtona dla układu dwóch równań nieliniowych startuje od pierwszego przybliżenia rozwiązania $(x_1^{(0)}, x_2^{(0)})$ i w pierwszym kroku prowadzi do

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} + \begin{bmatrix} h_1^{(0)} \\ h_2^{(0)} \end{bmatrix}.$$

W ogólnym przypadku dostajemy

$$\begin{bmatrix} x_1^{(n+1)} \\ x_2^{(n+1)} \end{bmatrix} = \begin{bmatrix} x_1^{(n)} \\ x_2^{(n)} \end{bmatrix} + \begin{bmatrix} h_1^{(n)} \\ h_2^{(n)} \end{bmatrix},$$

gdzie

$$\begin{bmatrix} h_1^{(n)} \\ h_2^{(n)} \end{bmatrix} = -J^{-1} \begin{bmatrix} f_1(x_1^{(n)}, x_2^{(n)}) \\ f_2(x_1^{(n)}, x_2^{(n)}) \end{bmatrix}.$$

Kryterium stopu

Obliczenia kończymy kiedy największa z różnic między kolejnymi iteracjami, dla każdej składowej, będzie odpowiednio mała, np. dla powyższego układu może to być

$$\max_{i=1,2} \left| x_i^{(n)} - x_i^{(n-1)} \right| \leq 10^{-12}.$$

Ćwiczenie 12. *Rozwiązać metodą Newtona układ równań*

$$\begin{cases} 2x_1 - x_2 + \frac{1}{9}e^{-x_1} = 1 \\ -x_1 + 2x_2 + \frac{1}{9}e^{-x_2} = 0, \end{cases}$$

rozpoczynając od punktu (1, 1), z dokładnością do 10^{-12} .

Ćwiczenie 13. *Rozwiązać metodą Newtona układ równań*

$$\begin{cases} xy = z^2 + 1 \\ xyz + y^2 = x^2 + 2 \\ e^x + z = e^y + 3, \end{cases}$$

dla punktu początkowego (1, 1, 1), z dokładnością do 10^{-10} .