

Spis treści

Przegląd podstawowych zagadnień związanych z wielomianami	1
Postawowe definicje	1
Postać iloczynowa wielomianu i dzielenie wielomianu	2
Pochodna wielomianu i jej obliczanie	2
Metoda Newtona oraz wielowymiarowa metoda Newtona	2
Opis klasycznej metody Newtona	2
Zastosowanie klasycznej metody Newtona do szukania zer wielomianu	3
Metoda Newtona dla wielu funkcji wielu zmiennych	3
Metoda Newtona dla funkcji zespolonych	4
Wybrane metody wyszukiwania miejsc zerowych wielomianu	5
Metoda Laguerre’a	5
Metoda Mullera	6
Metoda Bairstowa	6
Opis metody Bairstowa	6
Analiza teoretyczna metody Bairstowa	7
Przykłady rozbieżności metody Bairstowa	7
Testy numeryczne	7

Przegląd podstawowych zagadnień związanych z wielomianami

Postawowe definicje

Definicja 1. Wielomianem stopnia $n \in \mathbb{N}$ nad ciałem \mathbb{K} będziemy nazywać przekształcenie $\mathbb{K}^n \mapsto \mathbb{K}$ zadane wzorem $W(x) = a_0 + a_1x + \dots + a_nx^n$, gdzie a_i to pewne współczynniki z ciała \mathbb{K} .

Definicja 2. Niech W będzie pewnym wielomianem (nad ciałem \mathbb{K}). Liczbę a , taką że $W(a) = 0$, będziemy nazywać pierwiastkiem wielomianu.

Uwaga 1. Z faktu, że wielomian W ma współczynniki z ciała \mathbb{K} , nie wynika fakt, że jego pierwiastki również będą należeć do \mathbb{K} . Klasycznym przykładem jest wielomian $x^2 + 1$, który ma współczynniki rzeczywiste, a jego pierwiastkami są liczby zespolone.

Uwaga 2. Istnieją takie ciała \mathbb{K} , że dla dowolnego wielomianu stopnia większego od 0 wszystkie jego pierwiastki należą do \mathbb{K} . Ciała takie będziemy nazywać algebraicznie domkniętymi. Przykładem takiego ciała jest \mathbb{C} , czego nie będziemy dowodzić.

Podczas całego tego sprawozdania będziemy zajmować się następującym problemem:

Problem znajdowania miejsc zerowych wielomianu

Niech W będzie wielomianem. Celem jest znaleźć zbiór $\ker(W) = \{a \mid W(a) = 0\}$.

Powyższy problem, choć pozornie prosty, jest sformułowany bardzo ogólnie. Na potrzeby tej pracy od tej pory ograniczymy się tylko do \mathbb{R} oraz \mathbb{C} , choć nic nie staje na przeszkodzie by poeksperymentować z innymi ciałami. Aktualnie nie wiemy czy każdy wielomian ma pierwiastki, a jeśli ma to czy ich zbiór jest skończony. Nie znamy również żadnych metod rozwiązywania $W(x) = 0$. By lepiej zrozumieć podane zagadnienie przejdźmy przez

ciąg różnych definicji, algorytmów, twierdzeń i lematów związanych z wielomianami (warto je zrozumieć, gdyż kolejne rozdziały będą z nich korzystać).

Twierdzenie 1. *Każdy wielomian $W(x)$ nad \mathbb{C} stopnia $n \in \mathbb{N}_+$ ma co najmniej jeden pierwiastek.*

Dowód. To twierdzenie jest nazywane zasadniczym twierdzeniem algebry. Dowód [1] s. 105. \square

Wniosek 1. $|\ker(W)| \leq n$, gdzie n to stopień wielomianu W .

Postać iloczynowa wielomianu i dzielenie wielomianu

Definicja 3. *Wielomian $W(x)$ nazywamy podzielny przez wielomian $P(x)$, różny od wielomianu zerowego, wtedy i tylko wtedy, gdy istnieje taki wielomian $Q(x)$, że $W(x) = Q(x) \cdot P(x)$. Wielomian $Q(x)$ nazywamy ilorazem wielomianu $W(x)$ przez $P(x)$. Mówimy, że wielomian $P(x)$ jest dzielnikiem wielomianu $W(x)$.*

Definicja 4. *Dowolny wielomian $W(x)$ możemy zapisać jako $W(x) = P(x) \cdot Q(x) + R(x)$ dla pewnych wielomianów P, Q, R . Mówimy, że wielomian $W(x)$ jest podzielny przez $Q(x)$ jeżeli $R(x) = 0$.*

Twierdzenie 2. *Wielomian $W(x)$ jest podzielny przez wielomian $Q(x) = (x - a)$ wtedy i tylko wtedy, gdy $W(a) = 0$.*

Dowód. W [2] \square

Chcielibyśmy umieć w efektywny sposób realizować procedurę dzielenia wielomianu przez jednomiany postaci $x - a$. Służy do tego następujący algorytm:

1. $P(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$
2. Niech $\alpha = a_n$
3. Kolejno dla $k = n - 1, n - 2, \dots, 0$ wykonaj $\alpha := a_k + x\alpha$.
4. Wynik to $p(x) = \alpha$.

Dokładny opis metody oraz jej analizę możemy znaleźć w [3] (s 103).

Pochodna wielomianu i jej obliczanie

TUTAJ DODAC OPIS DOTYCZĄCY POCHODNYCH I OBLICZANIA ICH DLA WIELOMIANU.

Metoda Newtona oraz wielowymiarowa metoda Newtona

Opis klasycznej metody Newtona

Klasyczną metodą Newtona zastosowaną dla pewnego punktu startowego p oraz funkcji $f : \mathbb{R} \mapsto \mathbb{R}$ klasy C^1 nazywać będziemy metodę iteracyjną postaci:

$$x_n = \begin{cases} p, & \text{gdy } n = 0 \\ x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, & \text{w.p.p.} \end{cases}$$

Można pokazać, że x_n zbiega do pewnego pierwiastka funkcji f . Analizę klasycznej metody Newtona można znaleźć w [3] na stronach 71-81.

Zastosowanie klasycznej metody Newtona do szukania zer wielomianu

Jeśli mamy wielomian o współczynnikach i pierwiastkach rzeczywistych możemy policzyć jego pierwiastki za pomocą klasycznej metody Newtona. Podstawiamy za f z poprzedniego opisu nasz wielomian, a f' to jego pochodna. Po znalezieniu jednego pierwiastka (nazwijmy go a) dzielimy nasz wielomian przez $x - a$ i uruchamiamy program dla otrzymanego wielomianu. Proces kontynuujemy tak długo, aż dojdziemy do wielomianu o stopniu 0.

Uwaga 3. Wartość w punkcie wielomianu i jego pochodnej możemy wyznaczyć z pomocą schematu Hornera, który był omówiony wcześniej (w kodzie przykładowym skorzystaliśmy z funkcji bibliotecznych dla większej czytelności).

```
using Polynomials
# W – wielomian, n – stopień wielomianu, p – punkt startowy, eps – dokładność

function klasyczna_metoda_newtona(W, n, p, eps)
    dW = polyint(W)      # oblicza pochodną wielomianu
    x_n = p

    while abs(polyval(W, x_n)) >= eps      # dopóki błąd >= precyzja
        x_n = x_n - (polyval(W, x_n)/polyval(dW, x_n))
    end

    return x_n      # zwróć szukany pierwiastek
end
```

Uwaga 4. Powyższa metoda nie nadaje się do obliczania miejsc zerowych wielomianu, którego pierwiastki są zespolone (z powodu tego, że operujemy tutaj na tylko rzeczywistych przybliżeniach x_n).

Metoda Newtona dla wielu funkcji wielu zmiennych

Załóżmy, że mamy do rozwiązania układ równań:

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

gdzie $f_i \in \mathbb{R}^n \mapsto \mathbb{R}$ jest klasy C^1 .

Każdą z tych funkcji możemy rozpisać ze wzoru Taylora jako:

$$0 = f_i(x_1 + h_1, x_2 + h_2, \dots, x_n + h_n) \approx f_i(x_1, x_2, \dots, x_n) + \sum_{j=1}^n h_j \cdot \frac{\partial f_i}{\partial x_j}(x_1, x_2, \dots, x_n)$$

Powyższy układ możemy zapisać w postaci macierzowej:

$$\begin{pmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \dots \\ f_n(x_1, x_2, \dots, x_n) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_1, x_2, \dots, x_n) & \frac{\partial f_1}{\partial x_2}(x_1, x_2, \dots, x_n) & \dots & \frac{\partial f_1}{\partial x_n}(x_1, x_2, \dots, x_n) \\ \frac{\partial f_2}{\partial x_1}(x_1, x_2, \dots, x_n) & \frac{\partial f_2}{\partial x_2}(x_1, x_2, \dots, x_n) & \dots & \frac{\partial f_2}{\partial x_n}(x_1, x_2, \dots, x_n) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1}(x_1, x_2, \dots, x_n) & \frac{\partial f_n}{\partial x_2}(x_1, x_2, \dots, x_n) & \dots & \frac{\partial f_n}{\partial x_n}(x_1, x_2, \dots, x_n) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \dots \\ h_n \end{pmatrix}$$

Aby nieco skrócić ten układ, będziemy go zapisywać jako $F(X) = -J \cdot H$. Jeśli macierz J jest nieosobliwa, to układ ma rozwiązanie w postaci:

$$-J^{-1} \cdot F(X) = H$$

Ostatecznie wzór Newtona dla układu funkcji wielu zmiennych możemy wzorem:

$$X_{k+1} = X_k + H_k = X_k - J^{-1}(X_k)F(X_k)$$

Metoda Newtona dla funkcji zespolonych

Lemat 1. *Dowolną funkcję analityczną $f : \mathbb{C} \mapsto \mathbb{C}$ możemy zapisać jako*

$$f(z) = f(x + yi) = P(x, y) + iQ(x, y),$$

gdzie $x, y \in \mathbb{R}$, $P(x, y) \in \mathbb{R}$, $Q(x, y) \in \mathbb{R}$

Przykład 1.

$$f(z) = z^3 - 2z = f(x + iy) = (x + iy)^3 - 2(x + iy) = (x^3 - 3xy^2 - 2x) + i(3x^2y - y^3 - 2y) = P(x, y) + iQ(x, y)$$

Niech $f(z) = P(x, y) + iQ(x, y)$. Równanie $f(z) = 0$ możemy sprowadzić do układu równań $Q(x, y) = 0$ i $P(x, y) = 0$. Taki układ równań rozwiązujemy za pomocą metody Newtona dla funkcji wielu zmiennych.

$$v_{n+1} = v_n - \frac{f(v_n)}{f'(v_n)}$$

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - J^{-1} \begin{pmatrix} P(x_n, y_n) \\ Q(x_n, y_n) \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} \frac{\partial P}{\partial x}(x_n, y_n) & \frac{\partial P}{\partial y}(x_n, y_n) \\ \frac{\partial Q}{\partial x}(x_n, y_n) & \frac{\partial Q}{\partial y}(x_n, y_n) \end{pmatrix}^{-1} \begin{pmatrix} P(x_n, y_n) \\ Q(x_n, y_n) \end{pmatrix}$$

Ponieważ wielomian jest funkcją holomorficzną, to zachodzi równanie Cauchy'ego-Riemanna:

$$\frac{\partial P}{\partial x} = \frac{\partial Q}{\partial y}, \quad -\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

Oznaczając $P = P(x_n, y_n)$, $Q = Q(x_n, y_n)$, $P_x = \frac{\partial P}{\partial x}(x_n, y_n)$, $Q_x = \frac{\partial Q}{\partial x}(x_n, y_n)$ oraz korzystając ze wzoru na macierz odwrotną możemy uprościć wzór na metodę Newtona do postaci:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n - \frac{PP_x + QQ_x}{P_x^2 + Q_x^2} \\ y_n - \frac{PP_y + QQ_y}{P_x^2 + Q_x^2} \end{pmatrix}$$

```
using MultiPoly
```

```
#### a - lista współczynników wielomianu (np. [1,2,3] reprezentuje wielomian 1 + 2x + 3x^2)
# Funkcja dzieląca wielomian przez pierwiastek używając schematu Hornera
```

```
function horner(a, x0)
```

```
    n = length(a) - 1
```

```
    b = Array{Complex{BigFloat}, n}
```

```
    b[n] = a[n+1]
```

```
    for _k in 0:(n-2)
```

```
        i = n - _k
```

```
        b[i-1] = a[i] + b[i]*x0
```

```
    end
```

```
    return b
```

```
end
```

```

# W – współczynniki wielomianu jw, n – stopień wielomianu, x0 + i y0 – punkt startowy, eps
function complex_newton(W, n::Int, x0::Float64, y0::Float64, eps::Float64)

    if n == 0
        return
    elseif n == 1
        w = Complex128(-W[1])/Complex128(W[2])
        @printf("%.16lf_+\\t_%.16lf_i\\n_", real(w), imag(w))
        return
    end

    x, y = generators(MPoly{Float64}, :x, :y) # zmienne w wielomianie
    p = zero(MPoly{Complex128})

    for i in 1:(n+1)
        p = p + W[i] * (x+y*im)^(i-1)
    end

    xn = x0; yn = y0; P = real(p); Q = imag(p-P)
    Px = diff(P, :x); Py = diff(P, :y)
    Qx = diff(Q, :x); Qy = diff(Q, :y)

    while( abs( evaluate(p, xn, yn) ) >= eps )
        eP = evaluate(P, xn, yn); eQ = evaluate(Q, xn, yn)
        ePx = evaluate(Px, xn, yn); eQy = evaluate(Qy, xn, yn)
        ePy = evaluate(Py, xn, yn); eQx = evaluate(Qx, xn, yn)
        xn = xn - (eP * ePx + eQ * eQx)/(ePx^2 + eQx^2)
        yn = yn - (eP * ePy + eQ * eQy)/(ePx^2 + eQx^2)
    end

    @printf("%.16lf_+\\t_%.16lf_i\\n_", xn, yn)
    complex_newton(horner(W, complex(xn, yn)), n-1, x0, y0, eps)
end

```

Wybrane metody wyszukiwania miejsc zerowych wielomianu

Metoda Laguerre’a

Jedną z metod iteracyjnych wyszukiwania pierwiastków wielomianu używanych w nowoczesnych systemach informatycznych jest metoda Laguerre’a.

Niech $p(z)$ będzie wielomianem stopnia n , którego pierwiastki mamy znaleźć. Kolejne kroki w metodzie obliczamy za pomocą następujących wzorów:

$$A = \frac{-p'(z)}{p(z)}, \quad B = A^2 - \frac{p''(z)}{p(z)}, \quad C = \frac{A \pm \sqrt{(n-1)(nB - A^2)}}{n}, \quad z_{\text{nowe}} = z + \frac{1}{C}$$

Metoda Laguerre’a jest bardzo efektywnym algorytmem, ponieważ w okolicach pojedynczego pierwiastka wielomianu p jest zbieżna sześciennie. Dokładną analizę tej metody pozostawiamy czytelnikowi do przeczytania w [3] (s 112-116).

Metoda Mullera

Metoda Newtona jest modyfikacją metody stycznych. Zamiast przybliżać nasz wielomian f funkcją liniową będziemy go aproksymować funkcją kwadratową.

Rozważmy trzy punkty x_0, x_1, x_2 wraz z wartościami funkcji f w tych punktach. Przyjmujemy, że x_2 jest aktualnym przybliżeniem rozwiązania. Oznaczmy $z = x - x_2$, $h_0 = x_0 - x_2$, $h_1 = x_1 - x_2$.

Oznaczmy szukaną parabolę przez $g(z) = az^2 + bz + c$. Z definicji paraboli w punkcie $z = x_k$ dostajemy, że

$$2a = f''(x_k), \quad b = f'(x_k), \quad c = f(x_k),$$

co prowadzi do wzoru

$$x_{k+1} = x_k - \frac{2f(x_k)}{f'(x_k) + \operatorname{sgn}(f'(x_k)) \cdot \sqrt{(f'(x_k))^2 - 2f(x_k)f''(x_k)}}$$

Metoda Bairstowa

Opis metody Bairstowa

Kolejną i zarazem najważniejszą metodą, którą omówimy w sprawozdaniu będzie metoda Bairstowa. Wiemy, że nawet jeśli wielomian ma współczynniki rzeczywiste, to może mieć pierwiastki zespolone (np. $x^2 + 1$). Metoda Bairstowa pozwala na obliczenie wszystkich pierwiastków bez użycia arytmetyki zespolonej.

Lemat 2. *Jeżeli w jest pierwiastkiem nierzeczywistym wielomianu $p(z)$, a $p(z)$ jest wielomianem o współczynnikach rzeczywistych, to pierwiastkiem $p(z)$ jest również \bar{w} . Iloczyn $(x - w)(x - \bar{w})$ jest czynnikiem kwadratowym o współczynnikach rzeczywistych.*

Dowód. Dowód w [3] s. 108. □

Zauważmy, że pierwiastki zespolone możemy wyszukiwać parami. Zamiast wyszukiwać pierwiastki pojedynczo, będziemy wyszukiwać dwumianu postaci $z^2 - uz - v$.

Lemat 3. *Dowolny wielomian $p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_0$ możemy zapisać w postaci*

$$p(z) = (b_n z^{n-2} + b_{n-1} z^{n-3} + \dots + b_3 z + b_2) (z^2 - uz - v) + b_1 (z - u) + b_0$$

Wielomian $(b_n z^{n-2} + b_{n-1} z^{n-3} + \dots + b_3 z + b_2)$ będziemy dalej oznaczać jako $Q(z, u, v)$. Powyższe współczynniki możemy obliczać rekurencyjnie według wzorów:

$$b_{n+1} = b_{n+2} = 0, \quad b_k = ub_{k+1} + vb_{k+2} \quad (n \geq k \geq 0).$$

Dowód. Dowód w [3] s. 109. □

Chcemy by nasz wyjściowy wielomian był podzielny przez $z^2 - uz - v$. Zatem musi zachodzić $b_0 = b_1 = 0$. Potraktujmy podane współczynniki jako funkcje zmiennych u, v . Wtedy dostajemy do rozwiązania układ równań:

$$\begin{cases} b_0(u, v) = 0 \\ b_1(u, v) = 0 \end{cases}$$

Podany układ możemy rozwiązać przedstawioną wcześniej metodą Newtona dla wielu funkcji wielu zmiennych. Po znalezieniu współczynników u, v dzielimy wyjściowy wielomian przez otrzymany dwumian i kontynuujemy proces wyszukiwania pierwiastków dla mniejszego wielomianu (z uwzględnieniem tego, że przypadki dla wielomianu stopnia 0 i 1 traktujemy osobno).

Analiza teoretyczna metody Bairstowa

Lemat 4. *Metoda Bairstowa jest zbieżna lokalnie.*

Dowód. Wynika to bezpośrednio z tego, że metoda Newtona jest zbieżna lokalnie. \square

Głównym założeniem w lokalnej zbieżności metody Bairstowa jest to, że jacobian wyliczany przy metodzie Newtona się nie zeruje dla podanych wcześniej punktów startowych i kolejnych przybliżeń. Zastanówmy się w jaki sposób zerowanie się jacobianu zależy od punktów startowych oraz pierwiastków wielomianu.

Twierdzenie 3. *Niech u, v będą dowolnie wybranymi liczbami rzeczywistymi. Jakobian dla algorytmu Bairstowa jest macierzą odwracalną wtedy i tylko wtedy gdy $z^2 - uz - v$ oraz wielomian $Q(z, u, v)$ nie mają wspólnych pierwiastków. Rząd jacobianu jest jeden wtedy i tylko wtedy kiedy liczba wspólnych pierwiastków (z krotnościami) jest równa jeden. Jakobian się zeruje wtedy i tylko wtedy gdy $z^2 - uz - z$ dzieli $Q(z, u, v)$.*

Twierdzenie 4. *Założmy, że $P(z) = Q(z, u, v)(x^2 - u^*z - v^*)$ i założmy, że wyrażenia po prawej stronie nie mają wspólnego pierwiastka. Wtedy istnieje dodatnia liczba d taka że ciąg (u_k, v_k) generowany przez metodę Bairstowa jest zbieżny kwadratowo do (u^*, v^*) , gdzie $|u_0 - u^*| < d \wedge |v_0 - v^*| < d$.*

Dowód. Twierdzenia te zostały udowodnione przez autorów Tibora Fiałę oraz Annę Krebsz w 1987 roku. Kompletne dowody można przeczytać w [5]. \square

Analiza zbieżności oraz rozbieżności metody Bairstowa stanowiła podstawę do napisaniu kilku (choć niestety niewielu prac). Zainteresowanego czytelnika odsyłamy do [5], [6] oraz [7].

Przykłady rozbieżności metody Bairstowa

Testy numeryczne

Literatura

- [1] Leja, Franciszek. *Funkcje zespolone*, Warszawa : PWN, 1976
- [2] Aleksiej I., Kostrikin *Wstęp do algebry. Podstawy algebry*, Warszawa : PWN, 2008
- [3] David Kincaid, Ward Cheney *Analiza numeryczna*, Warszawa : WNT, 2006
- [4] Lily Yau, Adi Ben-Israel *The Newton and Halley Methods for Complex Roots*, The American Mathematical Monthly 105(1998), 806–818
- [5] Tibor Fiala, Anna Krebsz *On the Convergence and Divergence of Bairstow's Method*, Journal Numerische Mathematik, Volume 50 Issue 4, FEB. 1987, 477 - 482
- [6] Wolfgang Gabler *Invariances and convergence properties of Bairstow's method*, International Journal of Pure and Applied Mathematics Volume 27 No. 4 2006, 523-576
- [7] Sofo, Anthony and Glasson, Alan *Singularities in Bairstow's method*, Gazette of the Australian Mathematical Society, 37 (2). pp. 93-100. ISSN 0311-0729 1326-2297