

Metody numeryczne

Wykład nr 10

Miejsca zerowe wielomianów

Aneta Wróblewska

UMCS, Lublin

May 13, 2024

Poszukiwanie miejsc zerowych wielomianów jest fundamentalnym problemem w matematyce stosowanej, fizyce i inżynierii. Miejsce zerowe wielomianu to wartość, dla której wielomian przyjmuje wartość zero. Znalezienie tych miejsc ma kluczowe znaczenie w analizie równań różniczkowych, optymalizacji oraz w naukach inżynierskich.

Główne zagadnienia wykładu:

- Algorytm Hornera
- Wpływ zaburzeń współczynników
- Metoda Laguerre'a
- Obniżanie stopnia wielomianu i wygładzanie

Algorytm Hornera to efektywna metoda obliczania wartości wielomianu przy danej wartości zmiennej oraz wyznaczania jego pochodnych. Dzięki swojej strukturze pozwala na szybkie obliczenie wartości wielomianu, redukując liczbę mnożeń i dodawań, co jest szczególnie użyteczne w poszukiwaniu pierwiastków.

Wpływ zaburzeń współczynników

Wielomiany są bardzo wrażliwe na zaburzenia ich współczynników, co może prowadzić do znaczących zmian w lokalizacji ich miejsc zerowych. Zrozumienie tej wrażliwości jest kluczowe przy analizie błędów numerycznych oraz w przypadkach, gdy dane wejściowe są obarczone niepewnością.

Metoda Laguerre'a to jedna z zaawansowanych technik numerycznych służąca do znajdowania pierwiastków wielomianów. Jest to iteracyjna metoda, która może być stosowana do wszystkich typów wielomianów i jest znana z szybkiej zbieżności, szczególnie przy odpowiednim dobieraniu punktów startowych.

Obniżanie stopnia wielomianu i wygładzanie

Redukcja stopnia wielomianu oraz wygładzanie są technikami przetwarzania przedwstępnego, które mogą ułatwić znalezienie miejsc zerowych. Obniżanie stopnia polega na eliminowaniu terminów o niskim wpływie, podczas gdy wygładzanie może pomóc zredukować efekt oscylacji, szczególnie w wielomianach wysokiego stopnia.

Podstawowe Twierdzenie Algebry i schemat Hornera

Rozwiązanie równań wielomianowych

$$P_n(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = 0 \quad (1)$$

jest jednym z niewielu przypadków, w których interesuje nas znajomość wszystkich pierwiastków równania nieliniowego.

Podstawowe Twierdzenie Algebry

Wielomian stopnia n ma na płaszczyźnie zespolonej dokładnie n pierwiastków, przy czym pierwiastki wielokrotne liczy się z ich krotnościami.

Uwaga

W przypadku rzeczywistym wielomian stopnia n może nie mieć wcale pierwiastków, a jeśli ma, to jest ich co najwyżej n . Natomiast każdy wielomian rzeczywisty stopnia nieparzystego ma przynajmniej jeden pierwiastek rzeczywisty (wynika to z faktu, że granice niewłaściwe wielomianu rzeczywistego stopnia nieparzystego są różnych znaków, a także z faktu, że wielomian jako funkcja ciągła ma własność Darboux – a zatem musi przyjąć wartość pośrednią 0).

- **Charakter twierdzenia:** Podstawowe twierdzenie algebry nie jest konstruktywne — nie wskazuje sposobu poszukiwania pierwiastków — jednak dostarcza informacji o tym, czego szukać.
- **Przydatność teoretyczna:** Jest to częsty przykład, gdzie silny, chociaż niekonstruktywny wynik teoretyczny znacznie ułatwia stosowanie metod numerycznych.

- **Znaczenie:** Schemat Hornera to efektywny sposób obliczania wartości wielomianu oraz jego pochodnych w danym punkcie.
- **Koncepcja:** Polega na przekształceniu postaci wielomianu tak, aby minimalizować liczbę operacji arytmetycznych potrzebnych do jego wartości.

- Przekształcamy wielomian $P_n(z)$ w postać:

$$P_n(z) = (((a_n z + a_{n-1})z + a_{n-2})z + \cdots + a_1)z + a_0 \quad (2)$$

- Pamiętamy, że z każdym dodaniem kolejnego wyrazu redukujemy stopień wielomianu o jeden.
- Dzięki temu, potrzebujemy jedynie n mnożeń i n dodawań, aby obliczyć wartość wielomianu w punkcie z .

Algorithm 1: Algorytm schematu Hornera

Wejście: Wielomian

$$P_n(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = 0,$$

wartość z

Wyjście: Wartość wielomianu $P_n(z)$

```
1  $p \leftarrow a_n$ ;  
2 for  $i \leftarrow n - 1$  to 0 do  
3   |  $p \leftarrow p \cdot z + a_i$ ;  
4 end  
5 return  $p$ ;
```

Przykład: Rozważmy wielomian $P(z) = 3z^3 + 2z^2 - z + 5$ i chcemy obliczyć jego wartość dla $z = 2$.

Krok 1: $p \leftarrow 3$

Krok 2: $p \leftarrow p \cdot 2 + 2 = 3 \cdot 2 + 2 = 8$

Krok 3: $p \leftarrow p \cdot 2 - 1 = 8 \cdot 2 - 1 = 15$

Krok 4: $p \leftarrow p \cdot 2 + 5 = 15 \cdot 2 + 5 = 35$

Wartość wielomianu $P(z)$ dla $z = 2$ wynosi 35.

Schemat Hornera - podsumowanie

- **Zalety:** Schemat Hornera pozwala na efektywne obliczanie wartości wielomianu przy minimalnym zużyciu zasobów.
- **Zastosowania:** Jest szeroko stosowany w obliczeniach numerycznych oraz analizie numerycznej.
- **Wnioski:** Znając schemat Hornera, możemy zwiększyć wydajność naszych obliczeń, szczególnie w przypadku wielomianów o dużych stopniach.

Wpływ zaburzeń współczynników na miejsca zerowe wielomianów

Wpływ zaburzeń współczynników

W obliczeniach praktycznych współczynniki wielomianów, których pierwiastków poszukujemy, rzadko znamy w sposób dokładny. Najczęściej są one wynikiem jakichś poprzednich obliczeń, są zatem obarczone pewnymi błędami. Jaki jest wpływ błędów współczynników na wartości znalezionych numerycznie miejsc zerowych?

Wpływ zaburzeń współczynników (cd.)

Niech dokładny wielomian ma postać jak w równaniu 1 i niech z_0 będzie jego dokładnym pierwiastkiem. Przypuśćmy dalej, że dokładne wartości współczynników a_k nie są znane - zamiast tego znamy wartości przybliżone $\tilde{a}_k = a_k + \delta_k$, przy czym $\forall k : |\delta_k| \ll 1$. Poszukujemy pierwiastków wielomianu zaburzonego:

$$\tilde{P}_n(z) = \tilde{a}_n z^n + \tilde{a}_{n-1} z^{n-1} + \cdots + \tilde{a}_1 z + \tilde{a}_0 = 0 \quad (3)$$

Wpływ zaburzeń współczynników (cd.)

Spodziewamy się, że miejsce zerowe wielomianu 3 jest zaburzonym miejscem zerowym wielomianu 1: $\tilde{P}_n(\tilde{z}_0) = 0 \rightarrow \tilde{z}_0 = z_0 + \varepsilon$, $|\varepsilon| \ll 1$.

Mamy

$$\begin{aligned} 0 &= P_n(z_0) \\ &= a_n z_0^n + a_{n-1} z_0^{n-1} + \dots + a_1 z_0 + a_0 \\ &= (\tilde{a}_n - \delta_n)(\tilde{z}_0 - \varepsilon)^n + (\tilde{a}_{n-1} - \delta_{n-1})(\tilde{z}_0 - \varepsilon)^{n-1} + \dots \\ &\quad + (\tilde{a}_1 - \delta_1)(\tilde{z}_0 - \varepsilon) + (\tilde{a}_0 - \delta_0). \end{aligned} \tag{4}$$

Zauważmy, że

$$(\tilde{z}_0 - \varepsilon)^k = \sum_{l=0}^k \binom{k}{l} \tilde{z}_0^{k-l} (-1)^l \varepsilon^l \approx \tilde{z}_0^k - k \tilde{z}_0^{k-1} \varepsilon, \quad (5)$$

gdyż wyższe potęgi ε możemy zaniedbać. Zaniedbujemy również iloczyny $\delta_k \varepsilon$.

Zatem

$$\begin{aligned} 0 &= P_n(z_0) \\ &\approx \tilde{a}_n \tilde{z}_0^n - n \tilde{a}_n \tilde{z}_0^{n-1} \varepsilon + \delta_n \tilde{z}_0^n + \\ &\quad + \tilde{a}_{n-1} \tilde{z}_0^{n-1} - (n-1) \tilde{a}_{n-1} \tilde{z}_0^{n-2} \varepsilon + \delta_{n-1} \tilde{z}_0^{n-1} + \dots \\ &= \sum_{k=0}^n \tilde{a}_k \tilde{z}_0^k - \left(\sum_{k=1}^n k \tilde{a}_k \tilde{z}_0^{k-1} \right) \varepsilon + \sum_{k=0}^n \delta_k \tilde{z}_0^k. \end{aligned} \quad (6)$$

Wpływ zaburzeń współczynników (cd.)

Rozwijając wyrażenie, otrzymujemy:

$$\tilde{P}_n(\tilde{z}_0) = 0 + \left(\sum_{k=1}^n k \tilde{a}_k \tilde{z}_0^{k-1} \right) \varepsilon - \sum_{k=0}^n \delta_k \tilde{z}_0^k$$

Ostatecznie otrzymujemy następujące oszacowanie wpływu zaburzeń współczynników na zaburzenie miejsca zerowego wielomianu:

$$|\varepsilon| \approx \frac{|\sum_{k=0}^n \delta_k \tilde{z}_0^k|}{|\tilde{P}'_n(\tilde{z}_0)|} \quad (7)$$

Przykład Wilkinsona przedstawia wielomian:

$$W(z) = (z + 1)(z + 2) \dots (z + 20). \quad (8)$$

Jego miejsca zerowe to liczby całkowite ujemne: $-1, -2, \dots, -20$.

Założmy, że zaburzamy tylko jeden współczynnik w wielomianie

(8): $\delta_{19} = 2^{-23} \approx 10^{-7}$, $\delta_{k \neq 19} = 0$. Jak zmieni się położenie miejsca zerowego $z_0 = -20$?

Obliczamy $W'(-20) = -19!$. Oszacowanie (7) daje:

$$|\epsilon| \approx \frac{10^{-7} \cdot 20^{19}}{19!} \approx 4.4. \quad (9)$$

Zaburzenie miejsca zerowego jest siedem rzędów wielkości większe niż zaburzenie pojedynczego współczynnika! W rzeczywistości miejsca zerowe tak zaburzonego wielomianu stają się nawet zespolone. Zagadnienie znajdowania miejsc zerowych wielomianów może być źle uwarunkowane!

Zaburzenia wielokrotnych miejsc zerowych

Oszacowanie (7) załamuje się dla miejsc zerowych o krotności większej od jeden, co oznacza, że znikają także pochodne wielomianu. To sugeruje, że wielokrotne miejsce zerowe może zmienić się drastycznie po niewielkim zaburzeniu współczynników wielomianu.

Zaburzenia wielokrotnych miejsc zerowych (cd.)

Rozważmy przykład wielomianu:

$$\begin{aligned} Q(x) &= 39205740x^6 - 147747493x^5 + 173235338x^4 + 2869080x^3 - \\ &\quad 158495872x^2 + 118949888x - 28016640 \\ &= 17^3 \cdot 19 \cdot 20 \cdot 21 \left(x + \frac{20}{21}\right) \left(x - \frac{16}{17}\right)^3 \left(x - \frac{18}{19}\right) \left(x - \frac{19}{20}\right). \end{aligned}$$

(10)

Zaburzenia wielokrotnych miejsc zerowych (cd.)

Znalezienie miejsc zerowych tego wielomianu w postaci iloczynowej jest trywialne. Jednak numeryczne znalezienie miejsc zerowych tego samego wielomianu w postaci ogólnej może być bardzo trudne.

Zaburzenia wielokrotnych miejsc zerowych (cd.)

Mała zmiana w powyższym wielomianie , np. zwiększenie lub zmniejszenie wyrazu wolnego o 1, może spowodować przesunięcie potrójnych miejsc zerowych z osi rzeczywistej na płaszczyznę zespoloną.

Zaburzenia wielokrotnych miejsc zerowych (cd.)

W praktyce numerycznej, gdy otrzymujemy grupę miejsc zerowych bliskich "numerycznych miejsc zerowych", czasami trudno jest określić, czy są one naprawdę różne, czy też są one efektem skończonej dokładności, z jaką znamy współczynniki, i reprezentują "rozszczerzone" miejsce wielokrotne.

Poszukiwanie miejsc zerowych wielomianów

Co robić przy poszukiwaniu miejsc zerowych?

W kontekście poszukiwania miejsc zerowych wielomianów, niezbędne są dwie kwestie:

- Specjalistyczna metoda numeryczna do wyznaczania pierwiastków wielomianów. Taką metodą jest **metoda Laguerre'a**.
- Skuteczna strategia postępowania, która obejmuje:
 - Obniżanie stopnia wielomianu (tzw. deflacja) poprzez dzielenie wielomianu przez $(z - z_0)$, gdzie z_0 jest znalezionym pierwiastkiem.
 - Wygładzanie odkrytych miejsc zerowych za pomocą pierwotnego wielomianu, jeszcze przed jego deflacją.

Niech $P_n(z)$ będzie wielomianem stopnia n . **Metoda Laguerre'a** określona jest przez następującą iterację:

$$z_{i+1} = z_i - \frac{nP_n(z_i)}{P'_n(z_i) \pm \sqrt{(n-1)((n-1)(P'_n(z_i))^2 - nP_n(z_i)P''_n(z_i))}}, \quad (11)$$

gdzie znak w mianowniku wybierany jest tak, aby maksymalizować jego wartość bezwzględną.

Zbieżność metody Laguerre'a

- Jeśli wszystkie pierwiastki P_n są pojedyncze i rzeczywiste, metoda jest zbieżna sześciennie dla dowolnego rzeczywistego przybliżenia początkowego, tj. jeśli $|z_i - \bar{z}| < \epsilon \ll 1$, to $|z_{i+1} - \bar{z}| \sim \epsilon^3$, gdzie \bar{z} jest poszukiwanym pierwiastkiem.
- W ogólności metoda jest zbieżna sześciennie do wszystkich pojedynczych pierwiastków (rzeczywistych i zespolonych).

Ograniczenia i zalety metody Laguerre'a

- Metoda jest zbieżna liniowo do wielokrotnych pierwiastków.
- Metoda może nie zbiegać się w rzadkich przypadkach; stagnację można przełamać wykonując jeden-dwa kroki metody Newtona, a następnie wracając do metody Laguerre'a.
- Jest to metoda preferowana do wyszukiwania pierwiastków wielomianów, zarówno rzeczywistych, jak i zespolonych.

- Metoda Laguerre'a może być również stosowana do poszukiwania miejsc zerowych funkcji analitycznych rozwijalnych lokalnie w szereg potęgowy do rzędu n .
- Metoda Laguerre'a jest podobna do metody opartej o rozwinięcie w szereg Taylora do drugiego rzędu, ale uwzględnia stopień wielomianu, co czyni ją bardziej efektywną.

- Metoda wymaga obliczania drugiej pochodnej, co jest prostym zadaniem dla wielomianów.
- Nawet rozpoczynając od rzeczywistych punktów początkowych, metoda może prowadzić do zespolonych miejsc zerowych, co jest typowe dla wielomianów.

Deflacja - obniżanie stopnia wielomianu

Podczas numerycznego poszukiwania pierwiastków wielomianu może dojść do sytuacji, gdy różne próby zbiegają się do tego samego, już znalezionego pierwiastka. Aby tego unikać, stosujemy **deflację**, czyli technikę obniżania stopnia wielomianu poprzez faktoryzację:

$$P_n(z) = (z - z_1)P_{n-1}(z),$$

gdzie z_1 to pierwiastek wcześniej znaleziony dla wielomianu $P_n(z)$. Następnie szukamy pierwiastka dla nowo powstałego wielomianu $P_{n-1}(z)$.

Drobne zaburzenia współczynników wielomianu, które mogą powstać na skutek ograniczonej precyzji obliczeń, mogą znacząco wpłynąć na znalezione pierwiastki. Dlatego stosuje się **proces wygładzania**:

- Używamy pełnego, niepodzielonego wielomianu $P_n(z)$ do poprawy znalezionych miejsc zerowych.
- Założmy, że z_2 to przybliżone miejsce zerowe P_{n-1} . Używamy tego jako punktu startowego dla metody Laguerre'a.

Zastosowanie metody Laguerre'a

Zakładając, że numerycznie znalezione z_2 leży blisko prawdziwego pierwiastka, możemy spodziewać się szybkiej zbieżności przy użyciu metody Laguerre'a na wielomianie P_n . W rezultacie uzyskujemy dokładniejsze i wygładzone miejsce zerowe z_2 .

Po znalezieniu i wygładzeniu z_2 dla P_n , proces jest kontynuowany:

$$P_{n-1}(z) = (z - z_2)P_{n-2}(z)$$

co oznacza, że $P_n(z) = (z - z_1)(z - z_2)P_{n-2}(z)$. Następnie powtarzamy procedurę dla $P_{n-2}(z)$.

Deflacja wielomianu

Założmy, że z_0 jest pierwiastkiem wielomianu $P_n(z)$. Wtedy musi być spełniony związek:

$$\begin{aligned}(z - z_0) \cdot P_{n-1}(z) &= (z - z_0) (b_{n-1}z^{n-1} + b_{n-2}z^{n-2} + \dots + b_1z + b_0) \\ &= P_n(z).\end{aligned}\tag{12}$$

To równanie pozwala na obliczenie współczynników b_k rozwińmy to jako:

$$\begin{aligned}b_{n-1} &= a_n \\ -z_0 b_{n-1} + b_{n-2} &= a_{n-1} \\ &\dots \\ -z_0 b_1 + b_0 &= a_1 \\ -z_0 b_0 &= a_0\end{aligned}\tag{13}$$

Układ równań do deflacji

Wynikający układ równań liniowych z powyższego rozkładu można rozwiązać metodą podstawienia w przód:

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -z_0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & -z_0 & 1 \end{bmatrix} \begin{bmatrix} b_{n-1} \\ b_{n-2} \\ \vdots \\ b_1 \\ b_0 \end{bmatrix} = \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_2 \\ a_1 \end{bmatrix}. \quad (14)$$

Założmy, że znaleźliśmy już miejsca zerowe z_1, z_2, \dots, z_k wielomianu $P_n(z)$, używając deflacji do redukcji stopnia wielomianu, otrzymując nowy wielomian $P_{n-k}(z)$. Następnie:

- 1 Rozpoczynamy z dowolnym przybliżeniem i stosujemy metodę Laguerre'a do znalezienia kolejnego pierwiastka \tilde{z}_{k+1} wielomianu P_{n-k} .
- 2 W celu wygładzenia miejsce zerowego używamy \tilde{z}_{k+1} jako warunku początkowego do ponownego użycia metody Laguerre'a na pełnym wielomianie P_n , dla szybkiego uzyskania dokładniejszego wyniku z_{k+1} .
- 3 Wielomian $P_{n-k}(z)$ jest następnie faktoryzowany do $P_{n-k-1}(z)$, kontynuując proces aż do osiągnięcia wielomianu stopnia 2.

- Jeśli pierwotny wielomian $P_n(z)$ ma rzeczywiste współczynniki, jego pierwiastki są rzeczywiste lub tworzą sprzężone pary zespolone. Jeśli znajdziemy zespolony pierwiastek $z_k = x_k + iy_k$, to $z_{k+1} = x_k - iy_k$ również jest pierwiastkiem.
- Gdy pierwotny wielomian ma całkowite współczynniki, warto najpierw sprawdzić, czy posiada pierwiastki wymierne, zanim przystąpimy do numerycznych obliczeń.

- Rozważany wielomian: $P(x) = (x - 1)^3 = x^3 - 3x^2 + 3x - 1$
- Cel: Znalezienie wszystkich pierwiastków wielomianu.
- Metoda: Użycie metody Laguerre'a oraz techniki deflacji do systematycznego obniżania stopnia wielomianu i izolowania pierwiastków.

Metoda Laguerre'a

- Metoda Laguerre'a jest używana do efektywnego znajdowania pojedynczego pierwiastka wielomianu.
- Startujemy z przybliżonego pierwiastka, a metoda iteracyjnie poprawia to przybliżenie.
- Dla wielomianu $P(x)$, iteracja wygląda następująco:

$$z_{i+1} = z_i - \frac{nP(z_i)}{P'(z_i) \pm \sqrt{(n-1)[(n-1)(P'(z_i))^2 - nP(z_i)P''(z_i)]}}$$

- Wybór znaku w mianowniku zależy od tego, który daje większy moduł mianownika.

Pierwsza iteracja i deflacja

- Startowy punkt: $z_0 = 1.5$ (przybliżenie)
- Znaleziono pierwiastek: $z = 1.0$ (dla przykładu, w metodzie Laguerre'a ustalona została maksymalna liczba iteracji 100 w każdym kroku)
- Po znalezieniu pierwiastka $z_1 = 1.0$, stosujemy deflację:

$$P(x) = (x - 1)(x^2 - 2x + 1)$$

Druga iteracja i kolejna deflacja

- Teraz bierzemy $P(x) = x^2 - 2x + 1$ i szukamy kolejnych pierwiastków.
- Ponownie znajdujemy $z = 1.0$
- Deflacja daje nam:

$$P(x) = (x - 1)(x - 1)(x - 1)$$

Ostatnia iteracja

- Dla wielomianu $P(x) = x - 1$, pierwiastek jest oczywisty:
 $z = 1.0$
- Wszystkie pierwiastki wielomianu to $1.0, 1.0, 1.0$
- Proces jest zakończony, wszystkie pierwiastki zostały znalezione.

Podsumowanie

- Wszystkie pierwiastki wielomianu $(x - 1)^3$ zostały skutecznie znalezione za pomocą metody Laguerre'a i deflacji.
- Metoda pozwala na efektywne izolowanie i precyzyjne wyznaczanie pierwiastków, nawet jeśli są one wielokrotne.