

PRZYBLIŻONE ROZWIĄZYWANIE RÓWNAŃ I ICH UKŁADÓW

$$(1) \quad \begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0, \\ f_2(x_1, x_2, \dots, x_n) &= 0, \\ &\vdots \\ f_m(x_1, x_2, \dots, x_n) &= 0. \end{aligned}$$

Rozwiązaniem szczególnym układu równań (1) nazywamy jakikolwiek zespół liczb $\xi_1, \xi_2, \dots, \xi_n$ takich, że równania (1) są spełnione dla $x_1 = \xi_1, x_2 = \xi_2, \dots, x_n = \xi_n$.

Są takie zagadnienia, w których zależy nam jedynie na znalezieniu jakiegokolwiek rozwiązania szczególnego układu równań (1). Na ogół jednak celem naszym jest znalezienie rozwiązania pełnego tego układu.

¹⁾ W dziedzinie rzeczywistej będzie to zbiór wszystkich rozwiązań rzeczywistych, w dziedzinie zespolonej — zbiór wszystkich rozwiązań zespolonych. W rozdziale niniejszym ograniczymy się do dziedziny rzeczywistej.

rozwiązaniem pełnym. Postępowanie takie może oczywiście prowadzić do fałszywych wniosków. Aby można było uważać dany zbiór rozwiązań szczególnych za rozwiązanie pełne, konieczny jest dowód, że nie istnieją inne rozwiązania szczególne. Trudność dowodzenia nie może usprawiedliwiać braku dowodu.

Metody przybliżonego rozwiązywania układu równań (1) polegają zazwyczaj na podaniu algorytmu (tzn. postępowania rachunkowego) pozwalającego obliczać kolejno wyrazy ciągów przybliżeń

$$(2) \quad \begin{array}{l} \xi_{10}, \xi_{11}, \xi_{12}, \dots \\ \xi_{20}, \xi_{21}, \xi_{22}, \dots \\ \dots \dots \dots \\ \xi_{n0}, \xi_{n1}, \xi_{n2}, \dots \end{array}$$

takich, że

$$(3) \quad \lim_{j \rightarrow \infty} \xi_{ij} = \xi_i \quad (i=1, 2, \dots, n),$$

gdzie układ liczb $\xi_1, \xi_2, \dots, \xi_n$ jest jednym z rozwiązań szczególnych układu równań (1). Obliczając wartości $\xi_{1k}, \xi_{2k}, \dots, \xi_{nk}$ mamy zatem przybliżenie tego rozwiązania szczególnego. Nazywać je będziemy *k-tym przybliżeniem*. Liczbę *k* dobieramy tak, aby błędy przybliżenia były dostatecznie małe. Metodą obliczania przybliżeń (2) będą poświęcone wszystkie paragrafy niniejszego rozdziału począwszy od § 42.

Rozwiązywanie układu równań (1) rozpoczynamy jednak zazwyczaj od ustalenia liczby rozwiązań szczególnych tego układu i podania dla każdego rozwiązania szczególnego $\xi_1, \xi_2, \dots, \xi_n$ takich — możliwie małych — przedziałów $[a_i, b_i]$, że

$$(4) \quad a_i \leq \xi_i \leq b_i \quad \text{dla} \quad i=1, 2, \dots, n$$

i nierówności te nie są równocześnie spełnione dla żadnego innego rozwiązania szczególnego. Jest to często trudne do wykonania zadanie. W § 41 podamy niektóre z metod jego rozwiązywania, ale tylko dla równań z jedną niewiadomą.

Dobierając w (2) przybliżenia spełniające nierówności

$$a_i \leq \xi_{ij} \leq b_i \quad (i=1, 2, \dots, n; j=0, 1, \dots),$$

zapewniamy sobie, że rozwiązanie szczególne określone wzorami (3) jest tym, dla którego podaliśmy przedziały (4). Przez obliczenie przybliżeń kolejno dla wszystkich rozwiązań szczególnych otrzymujemy przybliżone rozwiązanie pełne układu (1).

Należy przy tym pamiętać o ocenie błędu uzyskiwanych przybliżeń, gdyż przybliżenie, o którego błędzie nie wiemy, jest dla nas

bezwartościowe. Niech $\bar{\xi}_1, \bar{\xi}_2, \dots, \bar{\xi}_n$ będą przybliżeniem rozwiązania szczególnego $\xi_1, \xi_2, \dots, \xi_n$ układu równań (1), niech

$$\delta_1 = \bar{\xi}_1 - \xi_1, \quad \delta_2 = \bar{\xi}_2 - \xi_2, \quad \dots, \quad \delta_n = \bar{\xi}_n - \xi_n$$

oraz

$$\varepsilon_1 = f_1(\bar{\xi}_1, \bar{\xi}_2, \dots, \bar{\xi}_n), \quad \varepsilon_2 = f_2(\bar{\xi}_1, \bar{\xi}_2, \dots, \bar{\xi}_n), \quad \dots, \quad \varepsilon_m = f_m(\bar{\xi}_1, \bar{\xi}_2, \dots, \bar{\xi}_n).$$

Liczby $\delta_1, \delta_2, \dots, \delta_n$ są zatem błędami przybliżeń $\bar{\xi}_1, \bar{\xi}_2, \dots, \bar{\xi}_n$, natomiast liczby $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$ są błędami, z jakimi te przybliżenia spełniają równania (1). O ile ocena błędów $\delta_1, \delta_2, \dots, \delta_n$ jest na ogół trudna, o tyle ocena błędów $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$ jest zwykle łatwa. Liczby $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$ mogą mieć znaczenie np. fizyczne i w niektórych zastosowaniach można się obejść bez oceny błędów $\delta_1, \delta_2, \dots, \delta_n$, gdyż wystarcza oszacowanie błędów $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$. Dostatecznie małym błędom $\delta_1, \delta_2, \dots, \delta_n$ odpowiadają małe błędy $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$, gdy funkcje $f_i(x_1, x_2, \dots, x_n)$ ($i=1, 2, \dots, m$) układu równań (1) są ciągłe. Dostatecznie małym błędom $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$ odpowiadają małe błędy $\delta_1, \delta_2, \dots, \delta_n$, jeśli założymy, że funkcje odwrotne do funkcji $f_i(x_1, x_2, \dots, x_n)$ są ciągłe.

Osobnej dyskusji wymagają błędy otrzymanych przybliżeń, gdy współczynniki układu równań (1) są przybliżone i obarczone błędami. Wpływ błędów współczynników na otrzymywane rozwiązania może być bardzo duży. Można konstruować przykłady, w których błędy rozwiązań spowodowane nieznacznymi błędami współczynników są dowolnie wielkie.

Osobnej dyskusji wymaga również sprawa dokładności obliczeń przybliżonych, jakie wykonujemy w trakcie rozwiązywania danego układu równań. Można podać przykłady układów kilku równań z kilku niewiadomymi takich, że w celu otrzymania przybliżeń pierwiastków, np. z dokładnością do 3 miejsc dziesiętnych, należy wykonywać rachunki z dokładnością do kilkudziesięciu (!) i więcej miejsc dziesiętnych.

Dyskusja dokładności otrzymywanych wyników opiera się na regułach przenoszenia błędów maksymalnych (por. Rozdział I).

W rozdziale niniejszym interesować nas będą jedynie metody poszukiwania rzeczywistych pierwiastków równań i ich układów, mimo że wiele z tych metod można stosować również w dziedzinie zespolonej, np. *regula falsi*, metoda Newtona i inne.

Niektóre z opisanych w rozdziale niniejszym metod sprowadzają rozwiązywanie układu równań do rozwiązywania układu równań liniowych. Jeżeli jest to układ o małej liczbie równań i niewiadomych, nie jest trudno go rozwiązać. Trudności pojawiają się dopiero, gdy liczba równań i niewiadomych jest duża. Ze względu na duże znaczenie praktyczne układów wielu równań liniowych z wieloma niewiadomymi, ich rozwiązywaniu poświęcimy cały rozdział następny.

A. Jedno równanie z jedną niewiadomą

§ 41. Metoda Sturma i metody pokrewne. Niech będzie dane równanie

$$(5) \quad f(x) = 0,$$

gdzie $f(x)$ jest funkcją ciągłą w danym przedziale $[a, \beta]$. Niech będzie też dany ciąg funkcji ciągłych w przedziale $[a, \beta]$

$$(6) \quad f(x) = f_0(x), f_1(x), f_2(x), \dots, f_m(x)$$

o następujących własnościach:

1° Jeżeli istnieje w danym przedziale taka liczba c , że $f_k(c) = 0$, gdzie $1 \leq k < m$, to $f_{k-1}(c)f_{k+1}(c) < 0$.

2° $f_m(x) \neq 0$ dla $a \leq x \leq \beta$.

3° Jeżeli funkcja $f(x)$ przechodząc przez punkt $x = d$ ($a \leq d \leq \beta$) zmienia znak (a więc $f(d) = 0$), to iloczyn $f(x)f_1(x)$ zmienia wtedy znak z ujemnego na dodatni.

Taki ciąg funkcji nazywamy *ciągłem Sturma*.

Oznaczmy symbolem $Z(t)$ liczbę kolejnych zmian znaku wyrazów tego ciągu w punkcie $x = t$ ($a \leq t \leq \beta$), z pominięciem wyrazów równych zeru w tym punkcie. Jeżeli np. dla $x = 2$ otrzymujemy z ciągu Sturma ciąg liczb

$$2, 0, -1, -3, 10, 15, -16, 3, 50, 113,$$

to piszemy $Z(2) = 4$, gdyż po pominięciu wyrazu drugiego jako równego zeru mamy kolejno cztery zmiany znaku: przy przejściu od $+2$ do -1 , od -3 do $+10$, od $+15$ do -16 i od -16 do $+3$.

Sformułujmy teraz następujące twierdzenie:

Twierdzenie Sturma. Jeżeli równanie (5) nie ma pierwiastków wielokrotnych, to liczba pierwiastków leżących w przedziale $a < x \leq b$, gdzie $a \geq a$ i $b \leq \beta$, jest równa $Z(a) - Z(b)$, to znaczy równa ubytkowi liczby zmian znaku w ciągu Sturma (6) przy przejściu od $x = a$ do $x = b$.

Dowód. Zauważmy nasamprzód, że na mocy własności 1°, jeżeli funkcja $f_k(x)$ ($k = 1, 2, \dots, m-1$) ma wartość zero w punkcie $x = t_0$, to funkcje $f_{k-1}(x)$ i $f_{k+1}(x)$ mają w tym punkcie wartości o różnych znakach. Ponieważ funkcje $f_{k-1}(x)$ i $f_{k+1}(x)$ są ciągłe, więc istnieje otoczenie punktu $x = t_0$, w którym mają one stałe, ale różne znaki. Jeżeli funkcja $f_k(x)$ zmienia znak, to jako funkcja ciągła musi przejść przez wartość zero. Ale w takim razie zmiana znaku $f_k(x)$ nie wpływa na liczbę $Z(t)$ w oto-

czeniu punktu t_0 , gdyż: albo $f_k(t)$ ma taki znak jak $f_{k-1}(t)$, a więc przeciwny niż $f_{k+1}(t)$, i w ciągu

$$(7) \quad f_{k-1}(t), f_k(t), f_{k+1}(t)$$

jest tylko jedna zmiana znaku, mianowicie przy przejściu od $f_k(t)$ do $f_{k+1}(t)$, albo $f_k(t)$ ma znak przeciwny niż $f_{k-1}(t)$, a więc taki, jak $f_{k+1}(t)$, i w ciągu (7) jest także tylko jedna zmiana znaku, mianowicie przy przejściu od $f_{k-1}(t)$ do $f_k(t)$, albo wreszcie $f_k(t) = 0$ i po pominięciu tej liczby jako zera w ciągu (7) mamy znowu tylko jedną zmianę znaku, mianowicie przy przejściu od $f_{k-1}(t)$ do $f_{k+1}(t)$.

Zatem zmiany znaków wartości funkcji $f_k(x)$ ($k=1, 2, \dots, m-1$) nie mają wpływu na liczbę $Z(a) - Z(b)$ w jakimkolwiek przedziale $a \leq x \leq b$. Ponieważ funkcja $f_m(x)$ na mocy własności 2° ma stały znak, więc na wartość $Z(a) - Z(b)$ mają wpływ tylko zmiany znaku funkcji $f(x)$. Funkcja $f(x)$ zmienia znak przechodząc przez każdy punkt $x = \xi$, w którym ma wartość zero, czyli przez punkt będący pierwiastkiem równania $f(x) = 0$, ponieważ założyliśmy, że nie ma ona pierwiastków wielokrotnych. Ale wtedy na mocy własności 3° ubywa zmiana znaku przy przejściu od $f(x)$ do $f_1(x)$ i liczba zmian znaków w ciągu Sturma maleje o jedną. Jeżeli liczby a i b nie są pierwiastkami funkcji $f(x)$, czyli pierwiastkami równania $f(x) = 0$, to oczywiście liczba $Z(a) - Z(b)$, jako ubytek liczby zmian znaków w ciągu Sturma przy przejściu od a do b , jest równa liczbie przejść funkcji $f(x)$ przez wartość zero w przedziale (a, b) , a więc równa liczbie pierwiastków równania $f(x) = 0$ w tym przedziale. Jeżeli liczba a jest pierwiastkiem tego równania, to w ciągu Sturma dla punktu a pomijamy pierwszy wyraz $f(a)$ jako równy zeru. Tym samym będzie $Z(t) = Z(a)$ dla $t \geq a$ w pewnym otoczeniu punktu $t = a$ i $Z(t) = Z(a) + 1$ dla $t < a$ w tymże otoczeniu. Zatem liczba $Z(a) - Z(b)$ nie zależy od tego, czy liczba a jest, czy nie jest pierwiastkiem równania $f(x) = 0$, czyli — inaczej mówiąc — jeżeli a jest pierwiastkiem równania $f(x) = 0$, to pierwiastek ten nie jest wliczany do liczby $Z(a) - Z(b)$ pierwiastków tego równania. Inaczej ma się rzecz w przypadku, gdy punkt $x = b$ jest pierwiastkiem równania $f(x) = 0$. Wtedy $Z(t) = Z(b)$ dla $t \geq b$ w pewnym otoczeniu punktu $t = b$ i $Z(t) = Z(b) + 1$ dla $t < b$ w tymże otoczeniu. Wynika stąd, że liczba $Z(a) - Z(b)$ uwzględnia ubytek wartości $Z(t)$ przy przejściu funkcji $f(x)$ przez wartość zero w punkcie $x = b$, czyli że pierwiastek $x = b$ równania $f(x) = 0$ jest wliczony do liczby $Z(a) - Z(b)$ pierwiastków tego równania.

Z powyższych rozważań wynika, że liczba $Z(a) - Z(b)$ jest liczbą pierwiastków równania $f(x) = 0$, leżących w przedziale $a < x \leq b$, co było do dowiedzenia.

Twierdzenie Sturma pozostaje prawdziwe, gdy jako przedział $a < x \leq b$ przyjąć przedział $-\infty < x < +\infty$, a w ciągu Sturma dopuścić wartości

niewłaściwe $+\infty$ albo $-\infty$. Wtedy $Z(-\infty)$ oznaczać będzie liczbę zmian znaków w ciągu

$$f(-\infty), f_1(-\infty), f_2(-\infty), \dots, f_m(-\infty),$$

którego wyrazami mogą być między innymi symbole $+\infty$ albo $-\infty$. Podobnie określamy $Z(+\infty)$. W ten sposób mamy następujący wniosek z twierdzenia Sturma:

Jeżeli równanie (5) nie ma pierwiastków wielokrotnych i $a = -\infty$, $\beta = +\infty$, to liczba jego pierwiastków jest równa $Z(-\infty) - Z(+\infty)$. (Mowa tu oczywiście tylko o pierwiastkach rzeczywistych, jak zresztą w całym niniejszym paragrafie).

Twierdzenie Sturma pozwala znaleźć przybliżone rozwiązanie pełne dowolnego równania (5), o ile to równanie nie ma pierwiastków wielokrotnych i potrafimy skonstruować ciąg Sturma (6). Liczba $Z(-\infty) - Z(+\infty)$ jest liczbą pierwiastków tego równania. Stosując twierdzenie Sturma do coraz to węższych przedziałów możemy znaleźć przybliżone wartości każdego z pierwiastków. Ten sposób postępowania będziemy nazywali *metodą Sturma*.

Twierdzenie Sturma jest bardzo mocnym środkiem pomagającym znajdować rozwiązanie pełne równania (5).

Pokażemy teraz, jak stosuje się metodę Sturma do rozwiązywania równań algebraicznych. Niech będzie dane równanie

$$(7) \quad w(x) = 0,$$

gdzie $w(x)$ jest wielomianem stopnia n ($n > 0$).

Ponieważ pochodna $w_1(x) = w'(x)$ jest wielomianem stopnia niższego niż $w(x)$, możemy podzielić $w(x)$ przez $w_1(x)$. Otrzymujemy jako iloraz wielomian $q_1(x)$ i jako resztę wielomian $r_1(x)$. Jeżeli wprowadzimy wielomian $w_2(x) = -r_1(x)$, to wykonane dzielenie można będzie zapisać następująco:

$$(8) \quad w(x) = q_1(x)w_1(x) - w_2(x).$$

Wielomian $w_2(x)$ jest niższego stopnia niż $w_1(x)$. Dzieląc $w_1(x)$ przez $w_2(x)$ analogicznie do dzielenia $w(x)$ przez $w_1(x)$, otrzymujemy równość

$$(9) \quad w_1(x) = q_2(x)w_2(x) - w_3(x),$$

gdzie $q_2(x)$ i $w_3(x)$ są wielomianami.

Postępując dalej analogicznie otrzymujemy ciąg wielomianów

$$(10) \quad w_0(x) = w(x), \quad w_1(x) = w'(x), \quad w_2(x), \quad w_3(x), \quad w_4(x), \dots$$

coraz to niższego stopnia. Z chwilą pojawienia się w tym ciągu stałej wszystkie następne wyrazy będą już stałe. Niech $w_m(x)$ oznacza pierwszą z kolei stałą.

Będziemy dalej zakładali, że

$$(11) \quad w_m(x) \neq 0,$$

wykażemy bowiem, że gdyby było $w_m(x) = 0$, to moglibyśmy sprowadzić rozwiązanie równania (7) do rozwiązania takich równań, dla których byłby spełniony warunek (11).

Gdyby mianowicie było $w_m(x) = 0$, to mielibyśmy

$$w_{m-2}(x) = q_{m-1}(x)w_{m-1}(x),$$

gdzie $w_{m-1}(x)$ byłby wielomianem co najmniej pierwszego stopnia. Wielomian $w_{m-2}(x)$ byłby, jak widzimy, podzielny przez $w_{m-1}(x)$. Ze wzoru

$$w_{m-3}(x) = q_{m-2}(x)w_{m-2}(x) - w_{m-1}(x)$$

wynika, że wielomian $w_{m-3}(x)$ byłby również podzielny przez $w_{m-1}(x)$. Rozumując dalej analogicznie dochodzimy do wniosku, że wielomian $w(x)$ byłby podzielny przez $w_{m-1}(x)$. Niech

$$v(x) = \frac{w(x)}{w_{m-1}(x)}.$$

Wtedy równanie (7) można by napisać w postaci

$$w_{m-1}(x) v(x) = 0.$$

Zatem jego rozwiązanie sprowadziłoby się do rozwiązania dwóch równań niższego stopnia

$$(12) \quad w_{m-1}(x) = 0,$$

$$(13) \quad v(x) = 0.$$

Gdyby dla któregośkolwiek z tych równań nie był spełniony warunek (11), to moglibyśmy jego rozwiązanie sprowadzić w analogiczny sposób do rozwiązania równań jeszcze niższego stopnia. Ponieważ zaś stopień równania (7) można obniżać co najwyżej n razy, więc postępowanie powyższe po skończonej ilości kroków doprowadzi nas zawsze do równań spełniających warunek (11).

Wykażemy teraz, że w przypadku gdy $w_m(x) \neq 0$ równanie (7) nie ma pierwiastków wielokrotnych.

Istotnie, gdyby równanie (7) miało pierwiastek k -krotny a ($k > 1$), to wtedy

$$w(x) = (x - a)^k u(x),$$

gdzie $u(x)$ byłby wielomianem stopnia $(n - k)$. Zatem

$$(14) \quad w'(x) = k(x - a)^{k-1} u(x) + (x - a)^k u'(x).$$

Widzimy, że oba wielomiany $w(x)$ i $w'(x)$ byłyby podzielne przez $(x-a)^{k-1}$. Wielomian $(x-a)^{k-1}$ nie redukowałby się do stałej, bo w przypadku pierwiastka wielokrotnego $k > 1$. Jeżeliby wielomiany $w(x)$ i $w_1(x) = w'(x)$ były podzielne przez $(x-a)^{k-1}$, to — jak wynika ze wzoru (8) — wielomian $w_2(x)$ byłby również przez ten czynnik podzielny. Gdyby zaś wielomiany $w_1(x)$ i $w_2(x)$ były podzielne przez $(x-a)^{k-1}$, to — jak wynika ze wzoru (9) — wielomian $w_3(x)$ byłby również przez ten czynnik podzielny. Rozumując dalej analogicznie dochodzimy do wniosku, że stała $w_m(x) \neq 0$ byłaby również podzielna przez $(x-a)^{k-1}$, co nie jest możliwe.

Wykażemy teraz, że ciąg

$$(15) \quad w_0(x) = w(x), \quad w_1(x) = w'(x), \quad w_2(x), \quad \dots, \quad w_m(x)$$

jest ciągiem Sturma w przedziale $-\infty < x < +\infty$.

Istotnie, jeżeli istnieje taka liczba c , że $w_k(c) = 0$ ($1 \leq k \leq m-1$), to $w_{k-1}(c) \neq 0$ i $w_{k+1}(c) \neq 0$. W przeciwnym bowiem razie na mocy (14) mielibyśmy $w_0(c) = w_1(c) = w_2(c) = \dots = w_m(c) = 0$. Ponieważ $w_m(x)$ jest stałą, więc mielibyśmy $w_m(x) = 0$ wbrew przyjętym założeniom. Skoro zatem $w_{k-1}(c) \neq 0$ i $w_{k+1}(c) \neq 0$, to na mocy (14)

$$w_{k-1}(c)w_{k+1}(c) < 0,$$

ponieważ $w_k(c) = 0$. W ten sposób stwierdziliśmy, że ciąg (15) ma własność 1° ciągu Sturma (6). Własność 2° jest spełniona, ponieważ stała $w_m(x) \neq 0$. Ponadto, jeżeli istnieje taka liczba d , że $w(d) = 0$ (tzn. liczba d jest pierwiastkiem równania (7)), to $w'(d) = w_1(d) \neq 0$, gdyż założyliśmy, że równanie (7) nie ma pierwiastków wielokrotnych. Wynika stąd, że funkcja $w(x)$ przechodząc przez punkt $x=d$ albo maleje, albo rośnie. W przypadku pierwszym $w'(x) = w_1(x) < 0$ w otoczeniu punktu $x=d$, natomiast $w(x) > 0$ dla $x < d$ i $w(x) < 0$ dla $x > d$. Zatem iloczyn $w(x)w'(x)$ zmienia znak przechodząc przez punkt $x=d$ z ujemnego na dodatni. W przypadku drugim jest $w'(x) > 0$ w otoczeniu punktu $x=d$, natomiast $w(x) < 0$ dla $x < d$ i $w(x) > 0$ dla $x > d$. Zatem i w tym przypadku iloczyn $w(x)w'(x)$ zmienia znak przechodząc przez punkt $x=d$ z ujemnego na dodatni. W ten sposób stwierdziliśmy, że ciąg (15) ma także własność 3° i jest ciągiem Sturma.

PRZYKŁAD 1. Rozwiązać metodą Sturma z dokładnością do 0,01 równanie

$$(16) \quad x^4 - 5x^3 + 2x - 11 = 0.$$

Mamy

$$w_0(x) = x^4 - 5x^3 + 2x - 11, \quad w_1(x) = w'_0(x) = 4x^3 - 15x^2 + 2.$$

Dzieląc $w_0(x)$ przez $w_1(x)$ otrzymujemy iloraz $\frac{1}{4}x - \frac{5}{16}$ i resztę

$$-\frac{75}{16}x^2 + \frac{3}{2}x - \frac{83}{8}.$$

Zatem

$$w_0(x) = \left(\frac{1}{4}x - \frac{5}{16}\right)(4x^3 - 15x^2 + 2) - \left(\frac{75}{16}x^2 - \frac{3}{2}x + \frac{83}{8}\right)$$

i stąd

$$w_2(x) = \frac{75}{16}x^2 - \frac{3}{2}x + \frac{83}{8}.$$

Dzieląc $w_1(x)$ przez $w_2(x)$ otrzymujemy analogicznie

$$w_1(x) = \left(\frac{64}{75}x - \frac{5488}{1875}\right)\left(\frac{75}{16}x^2 - \frac{3}{2}x + \frac{83}{8}\right) - \left(\frac{24832}{1875}x - \frac{60688}{1875}\right)$$

i stąd

$$w_3(x) = \frac{24832}{1875}x - \frac{60688}{1875}.$$

Dzieląc na koniec $w_2(x)$ przez $w_3(x)$ otrzymujemy

$$w_2(x) = \left(\frac{140625}{397312}x + \frac{463550625}{616628224}\right)\left(\frac{24832}{1875}x - \frac{60688}{1875}\right) + \frac{1337576875}{38539264}$$

i stąd

$$w_4(x) = -\frac{1337576875}{38539264}.$$

Ponieważ w ciągu Sturma interesują nas tylko znaki poszczególnych wyrazów, więc każdy z wyrazów tego ciągu możemy pomnożyć przez dowolną liczbę dodatnią. Mnożąc $w_2(x)$ przez 16, $w_3(x)$ przez $\frac{1875}{16}$ i $w_4(x)$

przez $\frac{38539264}{1337576875}$ otrzymujemy następujący ciąg Sturma:

$$\begin{aligned} w_0(x) &= x^4 - 5x^3 + 2x - 11, \\ w_1(x) &= 4x^3 - 15x^2 + 2, \\ w_2(x) &= 75x^2 - 24x + 166, \\ w_3(x) &= 1552x - 3793, \\ w_4(x) &= -1. \end{aligned}$$

(17)

Ponieważ

$$w_4(x) \neq 0,$$

więc równanie (16) nie ma pierwiastków wielokrotnych. Jeśli $Z(t)$ oznacza liczbę zmian znaku w ciągu Sturma (17) dla $x=t$, to na mocy twierdzenia Sturma liczba N pierwiastków rzeczywistych równania (16) jest równa

$$N = Z(-\infty) - Z(+\infty).$$

Dla $x = -\infty$ mamy w ciągu (17) wyrazy

$$+\infty, -\infty, +\infty, -\infty, -1$$

i trzy zmiany znaku, więc $Z(-\infty) = 3$.

Dla $x = +\infty$ mamy

$$+\infty, +\infty, +\infty, +\infty, -1$$

i jedną zmianę znaku, więc $Z(+\infty) = 1$. Stąd $N = 3 - 1 = 2$. Równanie (16) ma dwa pierwiastki rzeczywiste.

Dla $x = 0$ mamy w ciągu (17) wyrazy

$$-11, +2, +166, -3793, -1$$

i dwie zmiany znaku, więc $Z(0) = 2$. Na mocy twierdzenia Sturma liczba pierwiastków dodatnich jest równa

$$Z(0) - Z(+\infty) = 1,$$

a liczba pierwiastków niedodatnich równa

$$Z(-\infty) - Z(0) = 1.$$

Podstawiając coraz to nowe wartości na x w ciągu (17) i obserwując zmiany wartości funkcji $Z(x)$ możemy coraz dokładniej określać przedziały, w których leżą pierwiastki równania (16). Rachunki wygodnie jest ułożyć w tabelkę, w której piszemy tylko znaki wyrazów ciągu Sturma. Mamy

x	$w_0(x)$	$w_1(x)$	$w_2(x)$	$w_3(x)$	$w_4(x)$	$Z(x)$
-10	+	-	+	-	-	3
-5	+	-	+	-	-	3
-2	+	-	+	-	-	3
-1	-	-	+	-	-	2
-1,5	+	-	+	-	-	3
-1,3	+	-	+	-	-	3
-1,2	-	-	+	-	-	2
-1,25	-	-	+	-	-	2
-1,28	-	-	+	-	-	2

Stąd $x_1 \approx -1,29$.

Mamy dalej

x	$w_0(x)$	$w_1(x)$	$w_2(x)$	$w_3(x)$	$w_4(x)$	$Z(x)$
10	+	+	+	+	—	1
5	—	+	+	+	—	2
6	+	+	+	+	—	1
5,5	+	+	+	+	—	1
5,2	+	+	+	+	—	1
5,1	+	+	+	+	—	1
5,05	+	+	+	+	—	1
5,02	+	+	+	+	—	1

Stąd $x_2 \approx 5,01$.

Chociaż metodą Sturma można określić dokładnie liczbę pierwiastków równania (5) zawartych w danym przedziale $(a, b]$, to jednak rachunki do tego potrzebne są nieraz długie i mozolne.

W niektórych przypadkach można uprościć postępowanie korzystając z twierdzenia Budana-Fouriera. Jeżeli mianowicie funkcja $f(x)$ występująca w równaniu (5) ma w przedziale $[a, \beta]$ pochodne do n -tego rzędu włącznie i jej n -ta pochodna $f^{(n)}(x)$ jest w całym tym przedziale stałego znaku i różna od zera, to w miejsce ciągu Sturma (6) bierzemy następujący ciąg pochodnych:

$$(18) \quad f(x), f'(x), \dots, f^{(n)}(x).$$

Przeprowadzając dla tego ciągu rozważania analogicznie jak w dowodzie twierdzenia Sturma dochodzimy właśnie do następującego twierdzenia Budana-Fouriera:

Jeżeli $V(t)$ oznacza liczbę zmian znaku w ciągu (18) dla $x=t$, a N jest liczbą pierwiastków równania (5) w przedziale $a < x \leq b$, gdzie $a \geq a$ i $b \leq \beta$, z uwzględnieniem ich krotności, to

$$N = V(a) - V(b) - K,$$

gdzie K jest liczbą parzystą nieujemną.

Gdy $V(a) - V(b) = 0$ albo $V(a) - V(b) = 1$, to oczywiście

$$N = V(a) - V(b)$$

i twierdzenie Budana-Fouriera pozwala dokładnie określić liczbę N . Mogą jednak zachodzić przypadki, gdy nawet dla dowolnie małego przedziału $(a, b]$ liczba $V(a) - V(b)$ jest różna od 0 i od 1. Wtedy nie możemy z pomocą twierdzenia Budana-Fouriera rozstrzygnąć, czy i ile pierwiastków równania (5) zawiera się w tym przedziale, bo nie znamy liczby K . W takich przypadkach trzeba się uciec do innych metod, np. można zastosować metodę Sturma.

PRZYKŁAD 2. Zastosujemy twierdzenie Budana-Fouriera do rozwiązania równania (16). Ciąg (18) ma dla tego równania postać

$$(19) \quad \begin{aligned} f(x) &= x^4 - 5x^3 + 2x - 11, \\ f'(x) &= 4x^3 - 15x^2 + 2, \\ f''(x) &= 12x^2 - 30x, \\ f'''(x) &= 24x - 30, \\ f^{(4)}(x) &= 24. \end{aligned}$$

Dla $x = -\infty$ otrzymujemy w tym ciągu

$$+\infty, -\infty, +\infty, -\infty, +24$$

i stąd $V(-\infty) = 4$.

Dla $x = +\infty$ otrzymujemy

$$+\infty, +\infty, +\infty, +\infty, 24$$

i $V(+\infty) = 0$. Zatem liczba N pierwiastków równania (16) jest

$$N = 4 - 0 - K,$$

gdzie K jest liczbą parzystą nieujemną, a stąd równanie (16) może mieć 4 lub 2 pierwiastki rzeczywiste albo nie mieć ich wcale.

Dla $x = 0$ otrzymujemy z (19)

$$-11, +2, 0, -30, +24$$

i stąd $V(0) = 3$. Zatem równanie (16) ma $V(-\infty) - V(0) = 1$ pierwiastek niedodatni, a tym samym 3 lub 1 pierwiastek dodatni. Podstawiając w (19) coraz to nowe wartości na x dochodzimy — analogicznie jak w przypadku stosowania metody Sturm'a w przykładzie 1 — do wniosku, że jeden z pierwiastków równania (16) leży w przedziale $(-1,30, -1,28]$, a drugi w przedziale $(5,00, 5,02]$. Mamy stąd z dokładnością do 0,01: $x_1 \approx -1,29$ i $x_2 \approx 5,01$. Ponieważ jednak $V(5) = 1$ i $V(0) - V(5) = 2$, nie wiemy, czy w przedziale $(0,5]$ nie leżą jeszcze dwa pierwiastki rzeczywiste. Podstawiając w (19) $x = 1$ obliczamy, że $V(1) = 1$. W takim razie w przedziale $(1,5]$ na pewno pierwiastków nie ma, natomiast dla przedziału $(0,1]$ twierdzenie Budana-Fouriera w dalszym ciągu nie daje nam odpowiedzi. Dalsze zwężanie przedziału jest tu bezcelowe, należy uciec się do innych metod. W danym przypadku można bardzo łatwo rozstrzygnąć, czy w przedziale $0,1]$ są jeszcze pierwiastki równania (16), czy nie. Zauważamy mianowicie, że dla $0 < x \leq 1$ jest $x^4 - 5x^3 + 2x < 11$, a więc $f(x) < 0$ i w tym przedziale pierwiastków być nie może.

Z twierdzenia Budana-Fouriera wynika jako przypadek szczególny następujące twierdzenie:

TWIERDZENIE DESCARTES'A. *Jeżeli N jest liczbą pierwiastków dodatnich (z uwzględnieniem ich krotności) równania algebraicznego*

$$(20) \quad f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0 = 0 \quad (a_n \neq 0),$$

a L oznacza liczbę zmian znaku w ciągu współczynników

$$(21) \quad a_n, a_{n-1}, \dots, a_0,$$

to

$$N = L - K,$$

gdzie K jest pewną liczbą parzystą nieujemną.

Istotnie, wszystkie wyrazy $f^{(k)}(+\infty)$ ($k=0,1,\dots,n$) w ciągu (18) dla $x=+\infty$ mają ten sam znak co współczynnik a_n , zatem $V(+\infty)=0$. Natomiast dla $x=0$ wyraz $f^{(k)}(0)$ ma taki znak, jak współczynnik a_k (jeżeli $a_k=0$, to również $f^{(k)}(0)=0$), zatem $V(0)=L$, skąd $V(0)-V(+\infty)=L$ i na mocy twierdzenia Budana-Fouriera $N=L-K$, c.b.d.d.

Za pomocą twierdzenia Descartes'a można również oszacować liczbę pierwiastków ujemnych równania (20), ponieważ — jak łatwo zauważyć — jest ona równa liczbie pierwiastków dodatnich równania

$$a_n(-x)^n + a_{n-1}(-x)^{n-1} + \dots + a_1(-x) + a_0 = 0.$$

Na przykład dla równania (16) mamy następujący ciąg współczynników (21):

$$+1, -5, +2, -11.$$

W ciągu tym są trzy zmiany znaku, zatem na mocy twierdzenia Descartes'a równanie (16) ma 1 lub 3 pierwiastki dodatnie. Podstawiając w (16) $-x$ w miejsce x otrzymujemy równanie

$$(22) \quad x^4 + 5x^3 - 2x - 11 = 0,$$

dla którego mamy następujący ciąg współczynników

$$+1, +5, -2, -11.$$

W ciągu tym jest tylko jedna zmiana znaku, zatem na mocy twierdzenia Descartes'a równanie (22) ma 1 pierwiastek dodatni. Wynika stąd, że równanie (16) ma jeden pierwiastek ujemny.

Opisaliśmy tylko niektóre z najbardziej znanych metod ustalania liczby pierwiastków równania (5) leżących w określonym przedziale $(\alpha, \beta]$. Metodami tymi posługujemy się zazwyczaj jedynie dla ustalenia takich przedziałów $(a_1, b_1]$, $(a_2, b_2]$, ..., o których powiedzielibyśmy, że każdy zawiera dokładnie jeden z pierwiastków danego równania. Z chwilą ustalenia takich przedziałów szybciej na ogół dochodzimy do dokład-

niejszych przybliżeń szukanych pierwiastków stosując jedną z metod, którym będą poświęcone następne paragrafy, aniżeli stosując metody opisane w paragrafie niniejszym.

§ 42. Regula falsi. Dane jest równanie

$$(23) \quad f(x) = 0,$$

gdzie funkcja $f(x)$ jest w przedziale $[a, b]$ ciągła i albo ściśle rosnąca, albo ściśle malejąca. Zakładamy ponadto, że liczby $f(a)$ i $f(b)$ są przeciwnego znaku, tzn.

$$(24) \quad f(a)f(b) < 0.$$

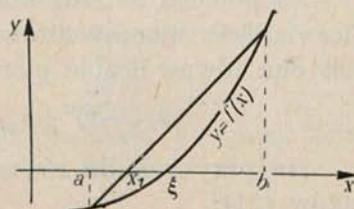
Z tych założeń wynika, że równanie (23) ma w przedziale (a, b) dokładnie jeden pierwiastek ξ . Podamy następującą prostą metodę obliczenia wartości przybliżonej x_1 tego pierwiastka: za x_1 przyjmujemy punkt dzielący przedział (a, b) w stosunku $|f(a)| : |f(b)|$, czyli na mocy (24) w stosunku $-f(a) : f(b)$. Wartość x_1 jest zatem współrzędną punktu przecięcia siecznej przechodzącej przez punkty $(a, f(a)), (b, f(b))$ z osią x (rys. 7).

Mamy zatem

$$\frac{x_1 - a}{b - x_1} = -\frac{f(a)}{f(b)},$$

skąd

$$(25) \quad x_1 = \frac{af(b) - bf(a)}{f(b) - f(a)} = a - \frac{(b-a)f(a)}{f(b) - f(a)}.$$



Rys. 7

Powyższy sposób obliczania nowego przybliżenia x_1 szukanego pierwiastka ξ , gdy znane są jego przybliżenia a i b , nosi nazwę *regula falsi* albo *metody siecznej*. Tą metodą można następująco otrzymać ciąg przybliżeń x_1, x_2, \dots . Jeżeli $f(x_1) = 0$, to oczywiście $x_1 = \xi$ i zadanie jest wykonane. Załóżmy zatem, że $f(x_1) \neq 0$. Jeżeli $f(x_1)f(a) < 0$, to obliczamy przybliżenie x_2 stosując wzór (25) do przedziału $[a, x_1]$, tzn.

$$x_2 = \frac{af(x_1) - x_1f(a)}{f(x_1) - f(a)} = x_1 - \frac{(x_1 - a)f(x_1)}{f(x_1) - f(a)}.$$

Jeżeli $f(x_1)f(a) > 0$, to wtedy $f(x_1)f(b) < 0$ i obliczamy x_2 stosując wzór (25) do przedziału $[x_1, b]$, tzn.

$$x_2 = \frac{x_1f(b) - bf(x_1)}{f(b) - f(x_1)} = x_1 - \frac{(x_1 - b)f(x_1)}{f(x_1) - f(b)}.$$

Przybliżenia x_3, x_4, \dots obliczamy analogicznie.

Jeżeli w ciągu x_1, x_2, \dots istnieje taki wyraz x_k , że $f(x_k) = 0$, to oczywiście $x_k = \xi$ i zadanie jest wykonane. Wykażemy, że jeżeli $f(x_i) \neq 0$ dla $i = 1, 2, \dots$, to ciąg przybliżeń x_1, x_2, \dots jest zbieżny do pierwiastka ξ .

Niech D_1, D_2, D_3, \dots będą kolejnymi przedziałami domkniętymi, do których stosujemy *regula falsi*, aby obliczyć x_1, x_2, x_3, \dots . Jest zatem

$$D_1 = [a, b], \quad D_2 = [a, x_1] \quad \text{albo} \quad D_2 = [x_1, b] \quad \text{itd.}$$

Każdy przedział D_k ($k = 2, 3, \dots$) jest zawarty w D_{k-1} , co zapisujemy następująco:

$$D_1 \supset D_2 \supset D_3 \supset \dots$$

Ciąg D_1, D_2, D_3, \dots jest więc ciągiem zstępującym przedziałów domkniętych.

Niech d_1, d_2, d_3, \dots będą długościami przedziałów D_1, D_2, D_3, \dots , tzn.

$$d_1 = b - a, \quad d_2 = x_1 - a \quad \text{albo} \quad d_2 = b - x_1 \quad \text{itd.}$$

Jest oczywiście

$$d_1 > d_2 > d_3 > \dots > 0,$$

więc ciąg d_1, d_2, d_3, \dots jest zbieżny, jako monotoniczny i ograniczony¹⁾.

Jeżeli

$$(26) \quad \lim_{n \rightarrow \infty} d_n = 0,$$

to istnieje dokładnie jeden punkt $x = \bar{x}$ wspólny wszystkim przedziałom D_1, D_2, D_3, \dots ²⁾. Ponieważ punkt $x = \xi$, gdzie ξ jest pierwiastkiem równania (23), leży wewnątrz każdego z przedziałów D_1, D_2, D_3, \dots , więc jest

$$\bar{x} = \xi.$$

Ponieważ dalej każda liczba x_k ($k = 1, 2, \dots$) jest jednym z krańców przedziału D_{k+1} , więc

$$|x_k - \xi| < d_{k+1}$$

i na mocy wzoru (26)

$$\lim_{n \rightarrow \infty} |x_n - \xi| = 0,$$

czyli

$$\lim_{n \rightarrow \infty} x_n = \xi,$$

co było do dowiedzenia.

¹⁾ Patrz np. F. Leja, *Rachunek różniczkowy i całkowy*, Warszawa 1954, str. 90.

²⁾ Ibidem, str. 90.

Jeżeli jednak

$$\lim_{n \rightarrow \infty} d_n = a \neq 0,$$

to

$$(27) \quad \lim_{n \rightarrow \infty} \frac{d_{n+1}}{d_n} = 1$$

i do każdej liczby q , takiej że $0 < q < 1$, można dobrać taką liczbę N , że

$$\frac{d_{n+1}}{d_n} > q$$

dla $n > N$. W szczególności na przykład

$$(28) \quad \frac{d_{n+1}}{d_n} > \frac{1}{2}$$

dla $n > N_0$, gdzie N_0 jest odpowiednio dobraną liczbą. Wynika już stąd że liczby $f(x_{k+1}), f(x_{k+2}), \dots$ dla $k > N_0$ będą miały jednakowe znaki. Niech mianowicie a_k i b_k będą krańcami przedziału D_k , tzn. $D_k = [a_k, b_k]$, i niech na przykład

$$|f(a_k)| < |f(b_k)|.$$

Stosując do przedziału D_k *regula falsi* otrzymujemy takie x_{k+1} , że

$$x_{k+1} - a_k < b_k - x_{k+1}$$

i na mocy (28) musi być

$$D_{k+1} = [x_{k+1}, b_k],$$

a zatem $f(x_{k+1})$ musi mieć znak przeciwny niż $f(b_k)$, czyli ten sam co $f(a_k)$, i ponadto z uwagi na monotoniczność funkcji $f(x)$

$$|f(x_{k+1})| < |f(a_k)| < |f(b_k)|.$$

Zatem następne przybliżenie x_{k+2} będzie takie, że

$$x_{k+2} - x_{k+1} < b_k - x_{k+2}$$

i na mocy (28) musi być

$$D_{k+2} = [x_{k+2}, b_k].$$

Rozumując dalej analogicznie dochodzimy do wniosku, że

$$D_n = [x_n, b_k]$$

dla $n > k > N_0$. Stąd na mocy (25)

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

czyli

$$\frac{d_n - d_{n+1}}{d_{n+1}} = \frac{1 - \frac{d_{n+1}}{d_n}}{\frac{d_{n+1}}{d_n}} = - \frac{f(x_n)}{f(b_k)}.$$

Z uwagi na (27) mamy stąd

$$(29) \quad \lim_{n \rightarrow \infty} f(x_n) = 0.$$

Z drugiej strony ciąg

$$x_{k+1}, x_{k+2}, \dots$$

jest rosnący i ograniczony, więc jest zbieżny. Niech

$$(30) \quad \lim_{n \rightarrow \infty} x_n = \bar{x}.$$

Ponieważ funkcja $f(x)$ jest ciągła z założenia, więc ze wzorów (29) i (30) wynika

$$\lim_{n \rightarrow \infty} f(x_n) = f(\bar{x}) = 0.$$

Zatem \bar{x} jest pierwiastkiem równania (23). Ponieważ to równanie może mieć tylko jeden pierwiastek ξ , więc

$$\bar{x} = \xi$$

i tym samym na mocy (30)

$$\lim_{n \rightarrow \infty} x_n = \xi,$$

co było do dowiedzenia.

W przypadku $|f(a_k)| > |f(b_k)|$ dowód przebiega analogicznie. W przypadku $|f(a_k)| = |f(b_k)|$ otrzymalibyśmy $d_{k+1} = d_k/2$, co przeczy założeniu (28)¹⁾.

Pozostaje sprawa oszacowania błędów otrzymywanych przybliżen x_1, x_2, \dots . Jeżeli — jak poprzednio —

$$D_n = [a_n, b_n], \quad n = 1, 2, 3, \dots$$

to mamy

$$(31) \quad |x_n - \xi| \leq b_n - a_n.$$

Jest to ogólne, ale najczęściej bardzo niedokładne oszacowanie.

¹⁾ Dowód powyższy można tak zmodyfikować, aby nie korzystać z monotoniczności funkcji. Wtedy liczba \bar{x} dana wzorem (30) będzie nadal pierwiastkiem równania (23), ale równanie to może mieć w przedziale (a, b) więcej pierwiastków. Można ponadto założenie ciągłości zastąpić założeniem ograniczoności funkcji $f(x)$ w przedziale $[a, b]$. Wtedy będzie również istniała granica $\lim_{n \rightarrow \infty} x_n = \bar{x}$, ale \bar{x} może już nie być pierwiastkiem równania (23).

Jeżeli funkcja $f(x)$ ma w przedziale $[a_n, b_n]$ pochodną $f'(x)$ taką, że

$$|f'(x)| > K > 0, \quad K = \text{const},$$

to na mocy twierdzenia o wartości średniej istnieje taki punkt $\bar{\xi}$, że $a_n < \bar{\xi} < b_n$ i

$$f(x_n) = f'(\bar{\xi})(x_n - \bar{\xi})$$

i mamy

$$(32) \quad |x_n - \bar{\xi}| < \frac{|f(x_n)|}{K}.$$

Z tego — na ogół bardzo dobrego — oszacowania możemy korzystać przy wszelkich metodach przybliżonych rozwiązywania równania $f(x) = 0$.

PRZYKŁAD 3. Rozwiązać równanie

$$(33) \quad x^3 - 2x - 5 = 0$$

z dokładnością do 0,000001.

Niech

$$f(x) = x^3 - 2x - 5.$$

Zauważmy, że

$$f(2) = -1, \quad f(3) = 16.$$

Zatem w przedziale (2,3) leży co najmniej jeden pierwiastek danego równania. Ponieważ $f'(x) = 3x^2 - 2 > 0$ dla $2 \leq x \leq 3$, więc funkcja $f(x)$ jest ściśle rosnąca w tym przedziale. Stosując zatem do równania (33) *regula falsi* dla przedziału [2,3] otrzymujemy ciąg przybliżeń zbieżny do pierwiastka ξ_1 zawartego w tym przedziale.

Stosując wzór (25) do przedziału [2,3] otrzymujemy

$$x_1 = 2 - \frac{1 \cdot (-1)}{16 - (-1)} \approx 2,06.$$

Obliczamy, że

$$f(2,06) \approx -0,378.$$

Zamiast stosować teraz *regula falsi* do przedziału [2,06,3] obliczamy, że

$$f(2,10) = 0,061$$

i stosujemy *regula falsi* do przedziału [2,06, 2,10]. Mamy w ten sposób nadal zapewnioną zbieżność ciągu otrzymywanych przybliżeń, ale szybkość zbliżania się do szukanego pierwiastka będzie większa.

Stosując zatem *regula falsi* do przedziału [2,06, 2,10] otrzymujemy

$$x_2 = 2,06 - \frac{0,04(-0,378)}{0,061 - (-0,378)} \approx 2,0944$$

i obliczamy, że

$$f(2,0944) \approx -0,00169, \quad f(2,0950) \approx 0,00501.$$

Stosując jeszcze raz wzór (25) otrzymujemy

$$x_3 = 2,0944 - \frac{0,0006(-0,00169)}{0,00501 - (-0,00169)} \approx 2,094551.$$

Oszacujemy błąd tego przybliżenia za pomocą wzoru (32). Mamy

$$f'(x) = 3x^2 - 2.$$

Dla przedziału $[2,0944, 2,0950]$ jest

$$|f'(x)| > 3 \cdot 2,0944^2 - 2 > 11,1.$$

Zatem

$$|x_3 - \xi_1| < \frac{|f(x_3)|}{11,1} < 0,00000051,$$

co już wypełnia warunki zadania.

Oszacowanie na podstawie wzoru (31) dałoby jedynie

$$|x_3 - \xi_1| < 0,0006$$

i musielibyśmy niepotrzebnie rachować dalej.

Pozostałe pierwiastki równania (33) znajdziemy rozwiązując równanie kwadratowe

$$x^2 + px + q = 0,$$

gdzie

$$x^2 + px + q = \frac{x^3 - 2x - 5}{x - \xi_1}.$$

Wykonując to dzielenie dochodzimy do równania

$$x^2 + 2,095x + 2,387 = 0,$$

które pierwiastków rzeczywistych nie posiada. Współczynniki w tym równaniu są przybliżone, ale zmiana ich wartości w granicach oszacowanego błędu nie może spowodować zmiany znaku wyróżnika.

Równanie (33) ma zatem tylko jeden pierwiastek rzeczywisty, którego wartość przybliżona jest

$$2,094551 \pm 0,00000051.$$

§ 43. Metoda Eulera. Dane jest równanie

$$(34) \quad f(x) = 0.$$

Niech to równanie ma w danym przedziale $[a, b]$ dokładnie jeden pierwiastek ξ , którego wartość przybliżoną x_0 znamy ($a \leq x_0 \leq b$). O funkcji $f(x)$

zakładamy, że ma w przedziale $[a, b]$ ciągle pochodne do rzędu n ($n \geq 1$) włącznie i że w przedziale $[a, b]$ jest $f'(x) \neq 0$. Zatem $f'(x)$ nie zmienia znaku, funkcja $f(x)$ jest monotoniczna i ma funkcję odwrotną

$$(35) \quad x = g(y).$$

Funkcję $x = g(y)$ rozwijamy według wzoru Taylora w otoczeniu punktu $y_0 = f(x_0)$

$$(36) \quad x = g(y_0) + \frac{g'(y_0)}{1!} (y - y_0) + \frac{g''(y_0)}{2!} (y - y_0)^2 + \dots + \frac{g^{(n)}(\eta)}{n!} (y - y_0)^n,$$

gdzie η jest pewną liczbą zawartą między y i y_0 .

Podstawiając w (36) $y = 0$ otrzymujemy $x = \xi$ i

$$(37) \quad \xi = g(y_0) - \frac{g'(y_0)}{1!} y_0 + \frac{g''(y_0)}{2!} y_0^2 - \dots + (-1)^n \frac{g^{(n)}(\eta)}{n!} y_0^n,$$

gdzie η leży między 0 i y_0 .

Pochodne funkcji $g(y)$ dają się wyrazić przez pochodne funkcji $f(x)$. Wprowadzając symbole $y'_0 = f'(x_0)$, $y''_0 = f''(x_0)$, ... mamy

$$\begin{aligned} g'(y_0) &= \frac{1}{y'_0}, \\ g''(y_0) &= -\frac{y''_0}{y'^3_0}, \\ g'''(y_0) &= (3y''^2_0 - y'_0 y'''_0) \frac{1}{y'^5_0}, \\ g^{(4)}(y_0) &= (-15y''^3_0 + 10y'_0 y''_0 y'''_0 - y'^2_0 y^{(4)}_0) \frac{1}{y'^7_0}, \\ &\dots \end{aligned}$$

Można wykazać (dowód pomijamy), że ogólnie

$$g^{(k)}(y_0) = (-1)^{k-1} \frac{P_k}{y'^{2k-1}_0},$$

gdzie P_k jest wyrażeniem dającym się obliczać rekurencyjnie ze wzoru

$$P_1 = 1, \quad P_{k+1} = (2k-1) y''_0 P_k - y'^2_0 (P_k)'. \quad .$$

Ponieważ ponadto $g(y_0) = x_0$, więc wzór (37) można napisać w postaci

$$(38) \quad \begin{aligned} \xi = x_0 - \frac{y_0}{y'_0} - \frac{1}{2!} \cdot \frac{y''_0}{y'_0} \left(\frac{y_0}{y'_0} \right)^2 - \frac{1}{3!} \left[3 \left(\frac{y''_0}{y'_0} \right)^2 - \frac{y'''_0}{y'_0} \right] \left(\frac{y_0}{y'_0} \right)^3 - \\ - \frac{1}{4!} \left[15 \left(\frac{y''_0}{y'_0} \right)^3 - 10 \frac{y''_0}{y'_0} \cdot \frac{y'''_0}{y'_0} + \frac{y^{(4)}_0}{y'_0} \right] \left(\frac{y_0}{y'_0} \right)^4 - \dots + R_n, \end{aligned}$$

gdzie $R_n = (-1)^n \frac{g^{(n)}(\eta)}{n!} y_0^n$, czyli

$$(39) \quad |R_n| \leq \frac{M_n}{n!} y_0^n,$$

$$a \quad M_n = \max_{a < x < b} |g^{(n)}(y)|.$$

Metoda Eulera polega na obliczaniu przybliżonej wartości pierwiastka ξ ze wzoru (38), w którym pomija się resztę R_n . Błąd tego przybliżenia można szacować ze wzoru (39) albo (32).

Ze wzoru (38) najwygodniej korzystać obliczając nasamprzód liczby

$$(40) \quad c_k = \frac{y_0^{(k)}}{y_0'} \quad (k=0, 2, 3, \dots).$$

Można go wtedy napisać następująco:

$$(41) \quad \xi = x_0 - c_0 - \frac{1}{2} c_2 c_0^2 - \frac{1}{6} (3c_2^2 - c_3) c_0^3 - \frac{1}{24} (15c_2^3 - 10c_2 c_3 + c_4) c_0^4 - \\ - \frac{1}{120} (105c_2^4 - 105c_2^2 c_3 + 15c_2 c_4 + 10c_3^2 - c_5) c_0^5 - \frac{1}{720} (945c_2^5 - \\ - 1260c_2^3 c_3 + 210c_2^2 c_4 + 280c_2 c_3^2 - 21c_2 c_5 - 35c_3 c_4 + c_6) c_0^6 - \dots + R_n.$$

PRZYKŁAD 4. Rozwiążemy zadanie z przykładu 3 stosując wzór Eulera.

Jak w przykładzie 3 zauważamy, że

$$f(2) = -1, \quad f(3) = 16.$$

Ponieważ

$$f'(x) = 3x^2 - 2,$$

a w przedziale $[2, 3]$ jest

$$3x^2 - 2 \geq 3 \cdot 2^2 - 2 > 0,$$

więc funkcja $f(x)$ jest w przedziale $[2, 3]$ rosnąca, ma pochodne dowolnego rzędu, funkcję odwrotną (35) i dokładnie jeden pierwiastek, dla którego mamy przybliżenie $x_0 = 2$.

Obliczamy kolejno

$$f'(x) = 3x^2 - 2, \quad f''(x) = 6x, \quad f'''(x) = 6, \quad f^{(4)}(x) = f^{(5)}(x) = \dots = 0.$$

Mamy zatem

$$y_0 = f(2) = -1, \quad y_0' = f'(2) = 10, \quad y_0'' = f''(2) = 12, \quad y_0''' = f'''(2) = 6, \\ y_0^{(4)} = y_0^{(5)} = \dots = 0,$$

a następnie ze wzoru (40)

$$c_0 = -0,1, \quad c_2 = 1,2, \quad c_3 = 0,6, \quad c_4 = c_5 = \dots = 0.$$

Podstawiając te liczby do wzoru (41), w którym uwzględniamy tylko wyrazy [do szóstej potęgi c_0 włącznie, otrzymujemy

$$\xi \approx 2 + 0,1 - 0,006 + 0,00062 - 0,000078 + 0,0000108 - 0,0000016 = 2,0945512.$$

Błąd otrzymanego przybliżenia można szacować według wzoru (39) albo (32). Wygodniejszy jest wzór (32).

Mamy w przedziale $[2,3]$

$$|f'(x)| = |3x^2 - 2| \geq 3 \cdot 2^2 - 2 = 10.$$

Ponadto $x_n = 2,0945512$ i $f(x_n) = -0,00000314$. Zatem ze wzoru (39)

$$|x_n - \xi| < 0,00000032 < 0,000001.$$

Błąd uzyskanego przybliżenia spełnia zatem warunki zadania.

Dyskusję istnienia innych pierwiastków przeprowadzamy tak samo jak w przykładzie 3.

§ 44. Metoda Newtona. Jeżeli we wzorze Eulera (38) przyjmiemy $n=2$, a zamiast symboli x_0, y_0, y'_0, R_n użyjemy symboli $x_m, f(x_m), f'(x_m), B_{m+1}$, to wzór ten przyjmie postać

$$(42) \quad \xi = x_m - \frac{f(x_m)}{f'(x_m)} + B_{m+1},$$

gdzie

$$(43) \quad B_{m+1} = -\frac{1}{2} \cdot \frac{f''(\bar{\xi})}{f'(\bar{\xi})} \left(\frac{f(x_m)}{f'(\bar{\xi})} \right)^2,$$

a $\bar{\xi}$ jest pewną liczbą zawartą między pierwiastkiem ξ a jego przybliżeniem x_m . Jeżeli we wzorze (42) opuścić resztę B_{m+1} , to zamiast pierwiastka ξ otrzymamy jego nowe przybliżenie x_{m+1} :

$$(44) \quad x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)}$$

z dokładnością określoną wzorem (43).

Rozpoczynając rachunek od danego przybliżenia x_0 możemy ze wzoru (44), dla $m=0,1,2,\dots$, obliczać coraz to nowe przybliżenia

$$(45) \quad x_0, x_1, x_2, \dots$$

Ten sposób postępowania nosi nazwę *metody kolejnych przybliżeń Newtona*. Ciągowi (45) odpowiada ciąg błędów B_1, B_2, B_3, \dots , określonych wzorem (43). Ze wzorów (42) i (44) otrzymujemy

$$(46) \quad B_{m+1} = \xi - x_{m+1}.$$