

RAFAŁ PALEJ*

SZEŚCIENNIE ZBIEŻNA METODA ROZWIĄZYWANIA UKŁADU NIELINIOWYCH RÓWNAŃ

CUBICALLY CONVERGENT METHOD FOR NONLINEAR EQUATION SYSTEMS

Streszczenie

W artykule przedstawiono metodę rozwiązywania układu nieliniowych równań algebraicznych. Zaprezentowana metoda jest uogólnieniem jednej z wersji metody Halleya. Metoda ta charakteryzuje się zbieżnością sześcienną i wymaga znajomości pochodnych pierwszego i drugiego rzędu, niezbędnych do wyznaczenia macierzy Jacobiego i Hessego.

Słowa kluczowe: układ nieliniowych równań algebraicznych, uogólniona metoda Halleya, zbieżność sześcienna

Abstract

The paper deals with the method for solving systems of nonlinear algebraic equations. The method presented in this paper is generalization about one of Halley's methods. This method is cubically convergent and it needs first and second derivatives in order to determine Jacobian and Hessians.

Keywords: system of nonlinear algebraic equations, generalized Halley's method, cubic convergence

* Dr hab. inż. Rafał Palej, prof. PK, Instytut Informatyki Stosowanej, Wydział Mechaniczny, Politechnika Krakowska.

1. Wstęp

Konieczność rozwiązywania układów nieliniowych równań algebraicznych pojawia się często podczas symulacji układów fizycznych. Waga problemu spowodowała duże zainteresowanie tą problematyką osób zajmujących się metodami numerycznymi. Chociaż historia rozwoju metod rozwiązywania równań i układów równań nieliniowych sięga czasów Newtona, to w kilku ostatnich dekadach można było odnotować dużo artykułów i monografii poświęconych tej tematyce. Zdecydowana większość metod opiera się na aproksymacji pochodnych [4], chociaż istnieje pogląd, że operowanie analitycznymi postaciami pochodnych ma swoje zalety [3]. Pomimo kwadratowej zbieżności metody Newtona, rekomendowanej ze względu na swą prostotę [1], pojawiło się wiele metod cechujących się wyższym stopniem zbieżności [5]. Podobnie wygląda sytuacja w odniesieniu do układu równań. Chociaż najprostszą i zarazem najpopularniejszą metodą rozwiązywania układów równań nieliniowych jest metoda Newtona-Raphsona, charakteryzująca się również zbieżnością kwadratową, to pojawiają się nowe metody cechujące się wyższym rzędem zbieżności [2]. W artykule przedstawiono metodę opierającą się na analitycznych wzorach opisujących pochodne pierwszego i drugiego rzędu, niezbędne do wyznaczenia macierzy Jacobiego i Hessego. Uwzględnienie we wzorze Taylora składników zawierających pochodne drugiego rzędu spowodowało, że przedstawiona w artykule metoda cechuje się zbieżnością sześcienną.

2. Sformułowanie problemu

Weźmy pod uwagę układ n nieliniowych równań algebraicznych ze względu na n niewiadomych, zapisany w postaci:

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}, \quad (1)$$

gdzie:

\mathbf{x} i \mathbf{f} – wektor niewiadomych i wektor lewych stron równań.

Związek pomiędzy wartością i -tej funkcji w punkcie $\mathbf{x}^{(2)}$, a wartościami tejże funkcji i jej pochodnych w punkcie $\mathbf{x}^{(1)}$ – oznaczającym pierwsze przybliżenie, dostarcza wzór Taylora:

$$f_i(\mathbf{x}^{(2)}) = f_i(\mathbf{x}^{(1)}) + \nabla f_i^T(\mathbf{x}^{(1)})\mathbf{h}^{(1)} + \frac{1}{2}\mathbf{h}^{(1)T}\mathbf{H}_i(\mathbf{x}^{(1)})\mathbf{h}^{(1)} + \dots, \quad (2)$$

gdzie $\mathbf{h}^{(1)} = \mathbf{x}^{(2)} - \mathbf{x}^{(1)}$, zaś gradient $\nabla f_i^T(\mathbf{x}^{(1)})$ i hesjan $\mathbf{H}_i(\mathbf{x}^{(1)})$ i -tej funkcji mają następującą budowę:

$$\nabla f_i^T(\mathbf{x}^{(1)}) = \left[\frac{\partial f_i}{\partial x_1} \Big|_{\mathbf{x}=\mathbf{x}^{(1)}} \quad \frac{\partial f_i}{\partial x_2} \Big|_{\mathbf{x}=\mathbf{x}^{(1)}} \quad \dots \quad \frac{\partial f_i}{\partial x_n} \Big|_{\mathbf{x}=\mathbf{x}^{(1)}} \right] \quad (3)$$

$$\begin{array}{cccc}
 \frac{\partial^2 f_i}{\partial x_1^2} & \frac{\partial^2 f_i}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f_i}{\partial x_1 \partial x_n} \\
 \frac{\partial^2 f_i}{\partial x_1 \partial x_2} & \frac{\partial^2 f_i}{\partial x_2^2} & \dots & \frac{\partial^2 f_i}{\partial x_2 \partial x_n} \\
 \vdots & \vdots & \ddots & \vdots \\
 \frac{\partial^2 f_i}{\partial x_1 \partial x_n} & \frac{\partial^2 f_i}{\partial x_2 \partial x_n} & \dots & \frac{\partial^2 f_i}{\partial x_n^2}
 \end{array} \quad (4)$$

Relację (2) można zapisać dla wszystkich funkcji równocześnie w formie:

$$\mathbf{f}(\mathbf{x}^{(2)}) = \mathbf{f}(\mathbf{x}^{(1)}) + \mathbf{J}(\mathbf{x}^{(1)})\mathbf{h}^{(1)} + \frac{1}{2} \begin{bmatrix} \mathbf{h}^{(1)T} \mathbf{H}_1(\mathbf{x}^{(1)}) \mathbf{h}^{(1)} \\ \mathbf{h}^{(1)T} \mathbf{H}_2(\mathbf{x}^{(1)}) \mathbf{h}^{(1)} \\ \vdots \\ \mathbf{h}^{(1)T} \mathbf{H}_n(\mathbf{x}^{(1)}) \mathbf{h}^{(1)} \end{bmatrix} + \dots \quad (5)$$

gdzie:

$\mathbf{J}(\mathbf{x}^{(1)})$ – macierz Jacobiego.

Poszukując wektora $\mathbf{h}^{(1)}$, dla którego $\mathbf{f}(\mathbf{x}^{(2)}) = 0$, otrzymamy nieliniowy układ równań na współrzędne wektora $\mathbf{h}^{(1)}$ w postaci:

$$\mathbf{f}(\mathbf{x}^{(1)}) + \mathbf{J}(\mathbf{x}^{(1)})\mathbf{h}^{(1)} + \frac{1}{2} \begin{bmatrix} \mathbf{h}^{(1)T} \mathbf{H}_1(\mathbf{x}^{(1)}) \mathbf{h}^{(1)} \\ \mathbf{h}^{(1)T} \mathbf{H}_2(\mathbf{x}^{(1)}) \mathbf{h}^{(1)} \\ \vdots \\ \mathbf{h}^{(1)T} \mathbf{H}_n(\mathbf{x}^{(1)}) \mathbf{h}^{(1)} \end{bmatrix} = 0 \quad (6)$$

Znajomość wektora $\mathbf{h}^{(1)}$ pozwoli określić wartość drugiego przybliżenia $\mathbf{x}^{(2)}$, a w konsekwencji wyprowadzić wzór iteracyjny.

3. Sposób rozwiązania

Dla $n = 1$ równanie (6) przedstawia równanie kwadratowe występujące w podejściu Halleya podczas rozwiązywania jednego równania nieliniowego. Aby uniknąć wyrażenia pierwiastkowego pojawiającego się w ścisłym rozwiązaniu tego równania, można wykorzystać rozwiązanie Newtona, otrzymując różne odmiany metody Halleya. Podejście to zostanie zastosowane przy wyprowadzaniu wzoru iteracyjnego służącego do wyznaczania przybliżonego rozwiązania układu równań (1).

W metodzie Newtona-Raphsona, opierającej się na wzorze Taylora uwzględniającym pochodne pierwszego rzędu, wektor $\mathbf{h}^{(1)}$ określony jest wzorem:

$$\mathbf{h}^{(1)} = -\mathbf{J}^{-1}(\mathbf{x}^{(1)})\mathbf{f}(\mathbf{x}^{(1)}) \quad (7)$$

Podstawiając wyrażenie (7) do trzeciego składnika równania (6), otrzymamy:

$$\mathbf{f}(\mathbf{x}^{(1)}) + \mathbf{J}(\mathbf{x}^{(1)})\mathbf{h}^{(1)} + \mathbf{r}(\mathbf{x}^{(1)}) = \mathbf{0} \quad (8)$$

gdzie wektor $\mathbf{r}(\mathbf{x}^{(1)})$ ma następującą budowę:

$$\mathbf{r}(\mathbf{x}^{(1)}) = \frac{1}{2} \begin{bmatrix} \mathbf{f}^T(\mathbf{x}^{(1)})\mathbf{J}^{-T}(\mathbf{x}^{(1)})\mathbf{H}_1(\mathbf{x}^{(1)})\mathbf{J}^{-1}(\mathbf{x}^{(1)})\mathbf{f}(\mathbf{x}^{(1)}) \\ \mathbf{f}^T(\mathbf{x}^{(1)})\mathbf{J}^{-T}(\mathbf{x}^{(1)})\mathbf{H}_2(\mathbf{x}^{(1)})\mathbf{J}^{-1}(\mathbf{x}^{(1)})\mathbf{f}(\mathbf{x}^{(1)}) \\ \vdots \\ \mathbf{f}^T(\mathbf{x}^{(1)})\mathbf{J}^{-T}(\mathbf{x}^{(1)})\mathbf{H}_n(\mathbf{x}^{(1)})\mathbf{J}^{-1}(\mathbf{x}^{(1)})\mathbf{f}(\mathbf{x}^{(1)}) \end{bmatrix} \quad (9)$$

Rozwiązując równanie (8) ze względu na $\mathbf{h}^{(1)}$ otrzymamy:

$$\mathbf{h}^{(1)} = -\mathbf{J}^{-1}(\mathbf{x}^{(1)})\mathbf{f}(\mathbf{x}^{(1)}) + \mathbf{r}(\mathbf{x}^{(1)}), \quad (10)$$

Drugie przybliżenie będzie określone wzorem:

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} - \mathbf{J}^{-1}(\mathbf{x}^{(1)})\mathbf{f}(\mathbf{x}^{(1)}) + \mathbf{r}(\mathbf{x}^{(1)}), \quad (11)$$

natomiast wzór iteracyjny przyjmie postać:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mathbf{J}^{-1}(\mathbf{x}^{(k)})\mathbf{f}(\mathbf{x}^{(k)}) + \mathbf{r}(\mathbf{x}^{(k)}), \quad k = 1, 2, \dots \quad (12)$$

Pomijając we wzorze (12) składnik $\mathbf{r}(\mathbf{x}^{(k)})$, otrzymamy wzór iteracyjny Newtona-Raphsona. Z kolei dla $n = 1$ otrzymamy ze wzoru (12) jedną z odmian metody Halleya w postaci [1]:

$$x_{k+1} = x_k - \frac{2f(x_1)f'(x_1) + f^2(x_1)f''(x_1)}{2f'(x_1)^3} \quad (13)$$

Można wykazać, że powyższy wzór iteracyjny cechuje się zbieżnością sześcienną dla x_1 dostatecznie bliskiego rozwiązania dokładnego.

4. Zbieżność metody

Wzór iteracyjny (12) można zapisać w postaci:

$$\mathbf{x}^{(k+1)} = \mathbf{g}(\mathbf{x}^{(k)}) \quad (14)$$

gdzie:

$$\mathbf{g}(\mathbf{x}^{(k)}) = \mathbf{x}^{(k)} - \mathbf{J}^{-1}(\mathbf{x}^{(k)})\mathbf{f}(\mathbf{x}^{(k)}) + \mathbf{r}(\mathbf{x}^{(k)}) \quad (15)$$

Wprowadzając pojęcie odległości k -tego przybliżenia od rozwiązania dokładnego \mathbf{x}^* wzorem:

$$\Delta\mathbf{x}^{(k)} = \mathbf{x}^* - \mathbf{x}^{(k)} \quad (16)$$

równanie (14) przyjmie postać:

$$\mathbf{x}^* - \Delta\mathbf{x}^{(k+1)} = \mathbf{g}(\mathbf{x}^* - \Delta\mathbf{x}^{(k)}) \quad (17)$$

Rozwijając $\mathbf{g}(\mathbf{x})$ w szereg Taylora wokół punktu \mathbf{x}^* z przyrostem $-\Delta\mathbf{x}^{(k)}$, otrzymamy:

$$\mathbf{g}(\mathbf{x}^* - \Delta\mathbf{x}^{(k)}) = \mathbf{g}(\mathbf{x}^*) - \mathbf{A}(\mathbf{x}^*)\Delta\mathbf{x}^{(k)} + \frac{1}{2} \begin{bmatrix} \Delta\mathbf{x}^{(k)T} \mathbf{B}_1(\mathbf{x}^*) \Delta\mathbf{x}^{(k)} \\ \Delta\mathbf{x}^{(k)T} \mathbf{B}_2(\mathbf{x}^*) \Delta\mathbf{x}^{(k)} \\ \vdots \\ \Delta\mathbf{x}^{(k)T} \mathbf{B}_n(\mathbf{x}^*) \Delta\mathbf{x}^{(k)} \end{bmatrix} - \dots \quad (18)$$

gdzie $\mathbf{A}(\mathbf{x}^*)$ i $\mathbf{B}_i(\mathbf{x}^*)$ oznaczają odpowiednio macierz Jacobiego i macierze Hessego utworzone z pochodnych elementów wektora $\mathbf{g}(\mathbf{x})$. Uwzględniając związek (18) w równaniu (17), otrzymamy:

$$\mathbf{x}^* - \Delta\mathbf{x}^{(k)} = \mathbf{g}(\mathbf{x}^*) - \mathbf{A}(\mathbf{x}^*)\Delta\mathbf{x}^{(k)} + \frac{1}{2} \begin{bmatrix} \Delta\mathbf{x}^{(k)T} \mathbf{B}_1(\mathbf{x}^*) \Delta\mathbf{x}^{(k)} \\ \Delta\mathbf{x}^{(k)T} \mathbf{B}_2(\mathbf{x}^*) \Delta\mathbf{x}^{(k)} \\ \vdots \\ \Delta\mathbf{x}^{(k)T} \mathbf{B}_n(\mathbf{x}^*) \Delta\mathbf{x}^{(k)} \end{bmatrix} - \dots \quad (19)$$

Ponieważ wektor \mathbf{x}^* jest rozwiązaniem równania $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, jest również rozwiązaniem równania $\mathbf{x} = \mathbf{g}(\mathbf{x})$, wobec czego równanie (19) określa związek zachodzący pomiędzy odległościami kolejnych przybliżeń w postaci:

$$\Delta\mathbf{x}^{(k+1)} = \mathbf{A}(\mathbf{x}^*)\Delta\mathbf{x}^{(k)} - \frac{1}{2} \begin{bmatrix} \Delta\mathbf{x}^{(k)T} \mathbf{B}_1 \Delta\mathbf{x}^{(k)} \\ \Delta\mathbf{x}^{(k)T} \mathbf{B}_2 \Delta\mathbf{x}^{(k)} \\ \vdots \\ \Delta\mathbf{x}^{(k)T} \mathbf{B}_n \Delta\mathbf{x}^{(k)} \end{bmatrix} + \dots \quad (20)$$

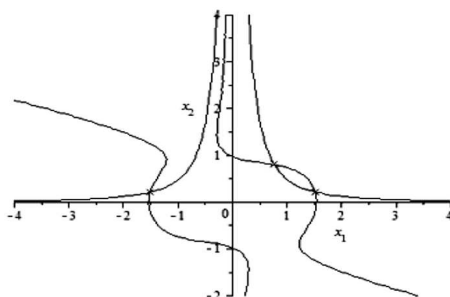
Wektor $\mathbf{g}(\mathbf{x})$, zdefiniowany wzorem (15), ma tę własność, że macierz Jacobiego $\mathbf{A}(\mathbf{x}^*)$ i macierze Hessego $\mathbf{B}_i(\mathbf{x}^*)$ zerują się. Oznacza to, że dla $\mathbf{x}^{(1)}$ dostatecznie bliskiego \mathbf{x}^* ciąg kolejnych przybliżeń, określony wzorem iteracyjnym (12), cechować się będzie zbieżnością sześciennej, uwidaczniającą się potrajaniem liczby cyfr znaczących w kolejnych iteracjach.

5. Przykład obliczeniowy

Przedstawiona w artykule metoda zostanie zilustrowana na przykładzie układu, dla którego wektor lewych stron równań ma postać:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_1 \sinh(x_1 x_2) - \frac{1}{2} \\ (x_1^2 + x_2^2)^2 - 2x_1^2 + 2x_1 x_2^5 - \frac{9}{10} \end{bmatrix} \quad (21)$$

Rozpatrywany układ równań ma 3 pierwiastki, których lokalizację przedstawia rys. 1.



Rys. 1. Lokalizacja pierwiastków rozpatrywanego układu równań
Fig. 1. Location of the roots of analyzed set of equations

Dla $\mathbf{x}^{(1)} = [0.8 \ 0.8]^T$ otrzymamy następujący ciąg przybliżeń jednego z rozwiązań.

Tabela 1

Cyfry znaczące w kolejnych przybliżeniach niewiadomej x_1

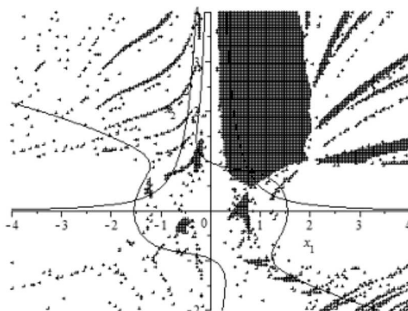
.8
.761425611366111550290464457701965529769055678717000233989116873878886397483396927
.761370793084825908919673403997264347473928506542185082038551103981570693053663583
.761370793084658464893797157379044840329561175400483940780993251209955224950011713
.761370793084658464893797157379044840322713393451290722806521706841037716765889666

Tabela 2

Cyfry znaczące w kolejnych przybliżeniach niewiadomej x_2

.8
.810149082552492346130457899443586892276442752449984741274062158713623861742503879
.810172721098292775151433878489822651633542621470508445682404980396632994094801492
.810172721098400086984127011343326549854429806188101679296084221760165113623239285
.810172721098400086984127011343326549859542144569948640916646907971687601290121281

Aby zaobserwować rząd zbieżności, obliczenia przeprowadzono z precyzją 85 cyfr. Interesujące z poznawczego punktu widzenia są obszary zbieżności do poszczególnych pierwiastków układu równań. Okazuje się, że nawet odległe punkty startowe mogą prowadzić do wyznaczenia przybliżenia poszukiwanego pierwiastka. Na rysunku 2 przedstawiono obszary zbieżności omawianej metody do pierwiastka $\mathbf{x}^* = [0.76... \ 0.81...]^T$.



Rys. 2. Obszary zbieżności do pierwiastka $\mathbf{x}^* = [0.76... \ 0.81...]^T$
Fig. 2. Regions of convergence to the root $\mathbf{x}^* = [0.76... \ 0.81...]^T$

Obszar zbieżności składa się z obszaru głównego – zawierającego poszukiwany pierwiastek – oraz szeregu podobszarów o zróżnicowanych rozmiarach. Pozostałe pierwiastki mają obszary zbieżności podobne w charakterze do obszaru przedstawionego na rys. 2.

5. Wnioski

Przedstawiona w artykule metoda charakteryzuje się prostym wzorem iteracyjnym. Zarówno macierz Jacobiego, jak i macierze Hessego można w łatwy sposób wyznaczyć, korzystając z dowolnego programu obliczeń symbolicznych. Z uwagi na złożoną budowę składnika $\mathbf{r}(\mathbf{x})$ metoda ta może być stosowana do układu równań o niewielkich rozmiarach. Obliczenia numeryczne pokazały, że obszar zbieżności zaprezentowanej w artykule metody niewiele różni się od obszaru zbieżności metody Newtona-Raphsona, cechującej się zbieżnością kwadratową.

Literatura

- [1] Dahlquist G., Björck Å., *Numerical Methods in Scientific Computing*, Vol. I, SIAM, Philadelphia 2008.
- [2] Homeier H.H.H., *A modified Newton method with cubic convergence: the multivariate case*, Journal of Computational and Applied Mathematics, Vol. 169, 2004.
- [3] Kelley C.T., *Solving Nonlinear Equations with Newton's Methods*, SIAM, Philadelphia 2003.
- [4] Ortega J.M., Rheinboldt W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM, Philadelphia 2000.
- [5] Wei L., Hong C., *A Unified Framework for the Construction of Higher-order Methods for Nonlinear Equations*, The Open Numerical Methods Journal, 2, 2010.