

TEORETYCZNE PODSTAWY METODY NEWTONA PRZYBLIŻONEGO ROZWIĄZYWANIA RÓWNAŃ NIELINIOWYCH

Aleksander Strasburger, Waclawa Tempczyk

Katedra Zastosowań Matematyki,

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

e-mails: aleksander_strasburger@sggw.pl; waclawa_tempczyk@sggw.pl

Streszczenie: Artykuł przedstawia podstawowe pojęcia teoretyczne i sformułowania leżące u podstaw wielowymiarowej metody Newtona konstruowania przybliżonych rozwiązań układu równań nieliniowych, a użytej w pracach [Strasburger i inni (2009), Strasburger i inni (2011)] dla modelowania pewnych aspektów teorii konsumpcji. Metoda ta korzysta ze stosunkowo prostych pojęć matematycznych i stanowi bardzo elastyczne narzędzie zarówno do rozważań teoretycznych jak i obliczeń numerycznych. Metoda ta z pewnością zasługuje na to, by być lepiej znaną w społeczności ekonomistów matematycznych.

Słowa kluczowe: wielowymiarowa metoda Newtona, metoda kolejnych przybliżeń, aproksymacja numeryczna.

WSTĘP

W opublikowanych niedawno dwóch pracach [Strasburger i in. 2009, Strasburger i in. 2011] jeden z autorów (A.S.) razem z A. Zembruskim przedstawili numeryczne symulacje pewnych zagadnień teorii konsumpcji, oparte na wykorzystaniu wielowymiarowej metody Newtona rozwiązywania równań nieliniowych. Jednakże w pracach tych nacisk był położony na procedurę obliczeniową, natomiast teoretyczne podstawy samej metody Newtona nie zostały tam omówione. Obecny artykuł ma charakter przeglądowy i jego zamierzeniem jest opisanie podstaw tej metody i przedstawienie warunków umożliwiających jej zastosowanie.

Metoda Newtona (nazywana czasami metodą Newtona-Raphsona) polega na konstruowaniu przybliżonych rozwiązań równania $f(x) = 0$, gdzie $f(x)$ jestadaną funkcją jednej zmiennej, algebraiczną lub nawet transcendentną. Metoda ta jest podawana w większości podręczników analizy matematycznej. Pomimo że jest ona względnie prostym zastosowaniem podstawowych idei analizy, jest ona jednocześnie metodą bardzo skuteczną i mającą wiele zastosowań. Mówiąc w skrócie, metoda polega na iteracyjnej konstrukcji ciągu (x_n) przybliżeń tego pierwiastka za pomocą następującej procedury: w każdym kroku konstrukcji funkcję $f(x)$ zastępujemy przez jej liniowe przybliżenie $l_n(x) = a_n x + b_n$ ze środkiem w punkcie x_n będącym aktualnym przybliżeniem, a jako następne przybliżenie x_{n+1} pierwiastka równania $f(x) = 0$, obieramy pierwiastek równania $l_n(x) = 0$.

Mówiąc dokładniej, przypomnijmy, że dla różniczkowalnej funkcji $f(x)$ jej liniowe przybliżenie ze środkiem w punkcie x_n dane jest wzorem

$$l_n(x) = f(x_n) + f'(x_n)(x - x_n) \quad (1)$$

gdzie f' oznacza pochodną funkcji f . Ciąg kolejnych przybliżeń rozwiązania zadany jest przez wybór początkowej wartości x_0 za pomocą rekurencyjnej formuły

$$x_{n+1} = x_n - \frac{1}{f'(x_n)} f(x_n) \quad (2)$$

Jest oczywiste, że funkcja f musi spełniać odpowiednie warunki, aby ciąg iteracji był określony i zbieżny do pierwiastka równania $f(x) = 0$. Możliwych jest wiele wyborów warunków zapewniających tę zbieżność — podamy sformułowanie zaczerpnięte z podręcznika [Walter 1992].

Twierdzenie 1 [Walter (1992)] *Załóżmy, że funkcja f jest określona i trzykrotnie różniczkowalna w sposób ciągły na przedziale $[a, b]$ osi rzeczywistej. Jeżeli pochodna $f'(x)$ nie znika na przedziale $[a, b]$ i pierwiastek ξ równania $f(x) = 0$ jest zawarty wewnątrz przedziału $[a, b]$, to ciąg iteracji (x_n) jest zbieżny do pierwiastka ξ niezależnie od wyboru punktu początkowego x_0 z pewnego otoczenia punktu ξ .*

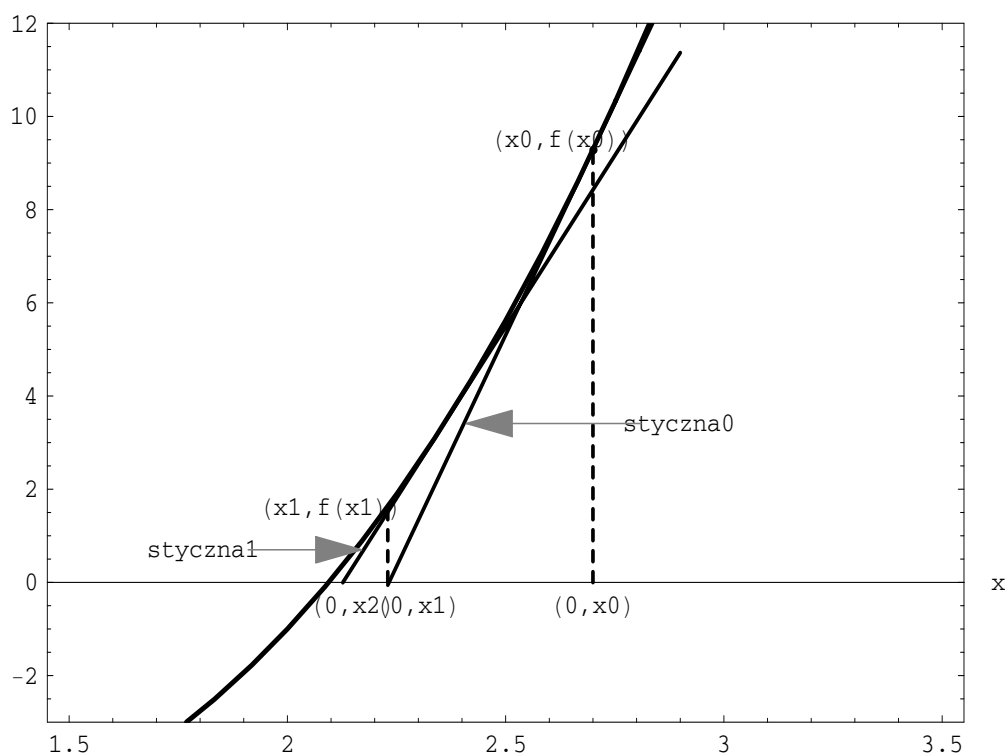
Kwestią szczególnej wagi, nie wspomnianą w tym sformułowaniu, jest wybór dopuszczalnej iteracji początkowej, gwarantujący zbieżność ciągu iteracji. Później odniesiemy się do tego problemu nieco szerzej.

Na zakończenie tych wstępnych rozważań dodajmy uwagę o szybkiej zbieżności metody Newtona w porównaniu z innymi metodami przybliżonymi. Jest to konsekwencją kwadratowego oszacowania dla błędu przybliżenia,

$$|\xi - x_{n+1}| \leq \text{const} \cdot |\xi - x_n|^2$$

które wynika po nieskomplikowanych rachunkach ze wzoru (2). Dokładność, o której mowa, można zaobserwować na załączonym rysunku, który przedstawia kilka początkowych przybliżeń pierwiastka równania $x^3 - 2x - 5 = 0$, które było rozpatrywane przez samego Newtona w nieopublikowanym rękopisie datowanym na 1669 r.

Rysunek 1 Ilustracja kolejnych przybliżeń pierwiastka równania $x^3 - 2x - 5 = 0$



Źródło: obliczenia własne w programie Mathematica ver. 5.0

WIELOWYMIAROWA METODA NEWTONA

Samo sformułowanie problemu w przypadku wielowymiarowym niewiele się różni od przypadku jednowymiarowego. Przy danym układzie k funkcji zależnych

od k zmiennych, $\phi_i(x_1, x_2, \dots, x_k)$, $i = 1, \dots, k$, poszukiwany jest taki punkt $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k)$, że każda z funkcji przyjmuje w tym punkcie wartość 0,

$$\phi_i(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k) = 0, \quad i = 1, 2, \dots, k \quad 3)$$

Użycie w tym miejscu odpowiedniej terminologii, bazującej na pojęciu funkcji przyjmującej wartości w przestrzeni \mathbb{R}^k , pozwala sformułować problem w sposób formalnie identyczny, co w przypadku jednowymiarowym. Wprowadźmy w tym celu odwzorowanie (funkcję o wartościach wektorowych) $\Phi: \mathbb{R}^k \rightarrow \mathbb{R}^k$ zdefiniowane wzorem

$$\Phi(x_1, x_2, \dots, x_k) = \begin{pmatrix} \phi_1(x_1, x_2, \dots, x_k) \\ \vdots \\ \phi_k(x_1, x_2, \dots, x_k) \end{pmatrix} = \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_k(x) \end{pmatrix},$$

gdzie $x = (x_1, x_2, \dots, x_k)$

Wówczas powiedzieć, że $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k)$ jest wspólnym miejscem zerowym układu równań (3) oznacza dokładnie tyle, co powiedzieć, że w tym punkcie funkcja Φ przyjmuje wartość 0 (oczywiście, tutaj $0 \in \mathbb{R}^k$).

Szukać pierwiastka czy punktu stałego?

Praktyczne zastosowania, szczególnie w problemach ekonomicznych, prowadzą częściej do zagadnienia wyznaczenia punktu stałego odwzorowania niż do problemu wyznaczenia pierwiastka równania. Z tego względu korzystniej zastąpić problem poszukiwania pierwiastka równania (układu równań) przez równoważny mu problem wyznaczania punktu stałego odpowiedniego odwzorowania, tym bardziej, że to podejście stawia do dyspozycji pokaźny zapas efektywnych narzędzi wypracowanych w teorii punktów stałych, w szczególności tak dopasowanych do stosowania metod numerycznych, jak metoda kolejnych przybliżeń.

W dalszej dyskusji będziemy używać następującej terminologii w odniesieniu do odwzorowania $\Phi: U \rightarrow U$:

Punktem stałym Φ nazywa się taki punkt $\xi \in U$, że $\Phi(\xi) = \xi$.

Dla dowolnej liczby naturalnej $n > 1$ n -krotne złożenie $\Phi^n = \underbrace{\Phi \circ \Phi \circ \dots \circ \Phi}_n$

odwzorowania Φ nazywamy iteracją stopnia n .

W teorii układów dynamicznych ciąg x^n o elementach określonych rekurencyjnie przez zadanie jakiegoś elementu początkowego $x^0 \in U$ i przyjęcie $x^n = \Phi(x^{n-1}) = \Phi^n(x^0)$ dla naturalnych $n \geq 1$, nazywa się orbitą punktu x^0 —

używana jest też nazwa ciąg iterowany o pierwszym elemencie (lub początkowej iteracji) x^0 .

Będziemy mówili, że odwzorowanie Φ jest kontrakcją (odwzorowaniem zbliżającym) ze współczynnikiem α , gdzie $1 > \alpha \geq 0$ jest stałą, jeśli dla wszystkich punktów $x, y \in U$ zachodzi nierówność

$$|\Phi(x) - \Phi(y)| \leq \alpha |x - y|$$

Łatwo zauważyć, że dowolna kontrakcja nie może mieć dwóch różnych punktów stałych, a znane twierdzenie Banacha o odwzorowaniach zbliżających podaje warunek dostateczny istnienia punktu stałego. Dla naszej analizy wystarczający jest następujący szczególny przypadek tego twierdzenia.

Twierdzenie 2. Niech $U \subset \mathbb{R}^k$ będzie zbiorem domkniętym oraz $\Phi: U \rightarrow U$ kontrakcją ze współczynnikiem α , $0 \leq \alpha < 1$. W zbiorze U istnieje dokładnie jeden punkt stały ξ odwzorowania Φ i jest on granicą orbity dowolnego punktu $x^0 \in U$, Inaczej mówiąc, dla dowolnego punktu $x^0 \in U$ ciąg iteracji o elemencie początkowym x^0 i wyrazach określonych zależnością

$$x^{n+1} = \Phi(x^n), \quad n = 0, 1, \dots$$

jest zbieżny do punktu stałego ξ odwzorowania Φ , $\lim_{n \rightarrow \infty} x^n = \xi$.

Szybkość zbiegania ciągu iteracji (x^n) do granicy można oszacować za pomocą nierówności

$$|x^n - \xi| \leq \frac{1}{1 - \alpha} |x^n - x^{n+1}| \leq \frac{(\alpha)^n}{1 - \alpha} |x^1 - x^0| \quad (4)$$

Wielowymiarowa metoda Newtona w pigułce

Przypomnijmy w tym miejscu elementy współczesnej notacji rachunku różniczkowego wielu zmiennych, która sprawia, że zapis metody Newtona w przypadku wielowymiarowym prawie nie odbiega od sposobu zapisu w przypadku jednowymiarowym.

Jeśli $x = (x_1, x_2, \dots, x_k)$ jest wektorem przestrzeni k -wymiarowej, (we wzorach zapisujemy wektory w postaci wektora kolumnowego) i $\Phi: U \rightarrow U$ jest odwzorowaniem o ciągłych pochodnych cząstkowych, to symbolem $\Phi'(x)$ oznaczana jest pochodna (macierz Jacobiego pochodnych cząstkowych odwzorowania Φ). Stosując dla pochodnych cząstkowych skrótowe oznaczenie

$$\partial_i \varphi_j(x) = \frac{\partial \varphi_j}{\partial x_i}(x), \quad \text{gdzie } i, j = 1, 2, \dots, k,$$

zapisujemy pochodną odwzorowania Φ w postaci następującej macierzy (macierzy Jacobiego)

$$\Phi'(x) = \begin{pmatrix} \partial_1 \varphi_1(x) & \dots & \partial_k \varphi_1(x) \\ \vdots & \vdots & \vdots \\ \partial_1 \varphi_k(x) & \dots & \partial_k \varphi_k(x) \end{pmatrix}$$

Dla oznaczenia iloczynu macierzy i wektora kolumnowego na ogół nie będziemy używać specjalnego symbolu, a jedynie w przypadkach wątpliwych będziemy rozdzielać je kropką jako znakiem mnożenia. Jeśli $\Phi'(x)$ jest macierzą odwracalną, tj. $\det(\Phi'(x)) \neq 0$, to macierz do niej odwrotną oznaczamy symbolem $(\Phi'(x))^{-1}$.

Założmy teraz, że pochodna $\Phi'(x)$ jest odwracalna na zbiorze U . Wówczas odwzorowanie

$$\Psi : x \mapsto x - (\Phi'(x))^{-1} \cdot \Phi(x) \quad (\text{IN})$$

jest dobrze określone, a równanie dla punktu stałego tego odwzorowania jest równoważne równaniu $\Phi(x) = 0$ dla pierwiastków Φ , czyli wspólnych miejsc zerowych układu równań (3). W ten sposób problem wyznaczenia pierwiastków układu nieliniowych równań (3) zostaje sprowadzony do problemu wyznaczenia punktów stałych odwzorowania Ψ określonego powyżej.

Możliwość zastosowania przedstawionego powyżej twierdzenia o punkcie stałym kontrakcji do rozważanego problemu wyznaczenia pierwiastka równania $\Phi(x) = 0$ zależy od dwóch czynników:

- wyboru niezmienniczej dziedziny dla odwzorowania Ψ ;
- wykazania dość przecież restryktywnego założenia o zbliżaniu dla Ψ .

Ten pierwszy wymaga określenia, dzięki umiejętnemu wyborowi lub odpowiedniej konstrukcji, domkniętego zbioru $U \subset \mathbb{R}^k$, takiego, że $\Psi(U) \subset U$, gdyż tylko wtedy można uruchomić procedurę rekurencyjną. Inaczej mówiąc, przy takim założeniu można w dowolnym kroku iteracji, wychodząc od skonstruowanego już n -tego elementu ciągu x^n , określić następny element ciągu wzorem $x^{n+1} = \Psi(x^n) = x^n - (\Phi'(x^n))^{-1} \cdot \Phi(x^n)$.

Ten drugi warunek zależy od analitycznych własności odwzorowania Φ , które są niejako dane z góry w sformułowaniu zadania i którymi nie daje się manipulować — trzeba je przyjąć „z dobrodziejstwem inwentarza”. W praktyce rzadko się zdarza, aby bez dodatkowych zabiegów warunki te były spełnione — oba na raz, lub nawet jeden z nich. Dlatego dostępne w literaturze przedmiotu rezultaty dotyczące wielowymiarowej metody Newtona mają przeważnie lokalny charakter i zapewniają zbieżność ciągu Newtona jedynie pod warunkiem odpowiedniego wyboru punktu początkowego ciągu iteracji. Dodatkowe utrudnienie, szczególnie ze względu na koszt numerycznych obliczeń, przynosi konieczność wyznaczania pochodnej $\Phi'(x^n)$ przy każdym kolejnym kroku

iteracji. Dlatego często stosuje się tak zwaną *modyfikowaną* metodę Newtona, w której zastępuje się we wzorze (IN) zależną od punktu pochodną $\Phi'(x)$ przez stałą macierz $\bar{J} = \Phi'(\bar{x})$ równą pochodnej obliczonej w odpowiednio dobranym punkcie \bar{x} . Przytoczymy tu sformułowanie oparte na dyskusji z monografii [Walter 1992] i [Strang 1986].

Twierdzenie 3. Niech $\bar{x} \in U$ będzie tak dobranym punktem, że macierz $\bar{J} = \Phi'(\bar{x})$ jest odwracalna. Niech dalej odwzorowanie $\bar{\Psi} : x \mapsto x - \bar{J} \cdot \Phi(x)$ będzie kontrakcją ze współczynnikiem $\alpha = \frac{1}{2}$ w otwartej kuli $K_r(\bar{x}) \in U$ o środku w \bar{x} i promieniu r . Jeśli $|\bar{J} \cdot \Phi(\bar{x})| < \frac{1}{2}r$, to odwzorowanie Φ ma w kuli $K_r(\bar{x})$ dokładnie jeden pierwiastek ξ . Ciąg iteracyjny Newtona $x^{n+1} = \bar{\Psi}(x^n) = x^n - \bar{J} \cdot \Phi(x^n)$, $n = 0, 1, 2, \dots$ jest dobrze określony dla dowolnego punktu początkowego $x^0 \in K_r(\bar{x})$ i zbiega do pierwiastka ξ równania $\Phi(x) = 0$, $\lim_{n \rightarrow \infty} x^n = \xi$.

BIBLIOGRAFIA

- Dudek H. (2011) Skale ekwiwalentności – estymacja na podstawie kompletnych modeli popytu, Wydawnictwo SGGW, Warszawa.
- Kelley C. T. (2003) Solving nonlinear equations with Newton's method, SIAM, Philadelphia.
- Krantz S. G., Parks H. R. (2002) The Implicit Function Theorem: History, Theory, and Applications, Birkhäuser, Boston.
- Panek E. (2000) Ekonomia matematyczna, Akademia Ekonomiczna w Poznaniu, Poznań.
- Strang, G. (1986) Introduction to Applied Mathematics, Wellesley-Cambridge Press, Wellesley.
- Strasburger A., Zembrzuski A. (2009) On application of Newton's method to solve optimization problems in the consumer and production theories, Polish J. of Environmental Studies, 2009, Vol. 18, 5B, pp. 198–202.
- Strasburger A., Zembrzuski A. (2011) On application of Newton's method to solve optimization problems in the consumer theory. Expansion's paths and Engel curves Metody Ilościowe w Badaniach Ekonomicznych, No 1, pp. 135-146.
- Walter, W., (1991), Analysis 1 & 2 (w jęz. niemieckim), Grundwissen Mathematik, Springer Verlag, Berlin.

**ON THEORETICAL FOUNDATIONS OF NEWTON'S METHOD
OF SOLVING NONLINEAR EQUATIONS**

Abstract: The paper surveys basic theoretical concepts and formulations underlying the multidimensional Newton's method of constructing approximate solutions to systems of nonlinear equations., which was used in the papers [Strasburger et al. (2009), Strasburger et al. (2011)] for modeling certain aspects of consumer's theory. This method is based on relatively simple mathematical concepts and is a very flexible tool for both theoretical arguments as well as for numerical computations. It certainly deserves to be better known in the community of mathematical economists.

Key words: multidimensional Newton's method, method of successive approximations, numerical approximations