Przez wieki z metodą Newtona

Piotr KRZYŻANOWSKI, Grzegorz ŁUKASZEWICZ**

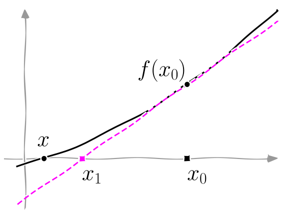
* Instytut Matematyki Stosowanej i Mechaniki, Uniwersytet Warszawski

Tytułowa metoda służy do wyznaczania przybliżenia miejsca zerowego x zadanej funkcji f, czyli innymi słowy – przybliżonego rozwiązywania równań postaci

$$f(x) = 0.$$

Główną ideę najprościej przedstawić na rysunku. Zaczynamy od jakiegoś sensownego przybliżenia, x_0, poszukiwanego miejsca zerowego x. Jak wiadomo, gładką krzywą można w okolicy x_0 dobrze aproksymować styczną do niej, wobec tego miejsce zerowe f powinno dawać się lepiej przybliżyć jako miejsce zerowe x_1 prostej stycznej do f w punkcie x_0 (zob. rysunek obok), a je przecież łatwo wyznaczyć:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$



Geometryczna interpretacja jednego kroku metody Newtona jako **metody stycznych**

Powyższy pomysł możemy powtarzać, dostając ciąg kolejnych przybliżeń x_0, x_1, x_2 itd., zadanych wzorem rekurencyjnym

(1) $$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

– i właśnie ten algorytm nazywamy **metodą Newtona**. Może więc nas zdziwić, że powyższy wzór (i to od razu z bardzo ważnym uogólnieniem na przypadek *układów* równań nieliniowych z wieloma niewiadomymi!) został po raz pierwszy podany przez Thomasa Simpsona w 1740 roku – a więc wtedy, gdy Isaac Newton od kilkunastu lat spoczywał w grobowcu w Opactwie Westminsterskim... Co więcej, na jej geometryczny aspekt, od którego tu zaczęliśmy, zwrócono uwagę nawet jeszcze później.

Za to *metodą Simpsona* nazywamy pewien sposób przybliżonego obliczania całek (choć Bonaventura Cavalieri wpadł na ten sam pomysł już w 1639 r.). Skądinąd jest to szczególny przypadek ogólnej metody aproksymacji całek, wcześniej odkrytej przez... Newtona.

Metoda Newtona, jakiej używał Newton

Dlaczego zatem „metoda Newtona”? W traktacie *De analysis per aequationes numero terminorum infinitas,* pochodzącym prawdopodobnie z 1669 roku, Newton przedstawia rozwiązywania równań wielomianowych, posługując się takim oto przykładem:

(2) $$x^3 - 2x - 5 = 0.$$

Wcześniej, zapewne około 1665 roku, do przybliżonego rozwiązywania podobnych równań wielomianowych Newton używał czegoś, co obecnie nazywamy **metodą siecznych** i traktujemy jako przybliżoną wersję metody... stycznych.

Ponieważ na pierwszy rzut oka widać, że pierwiastkiem tego równania jest liczba około 2, to zapisując go w postaci $x = 2 + p$ i podstawiając do równania (2), dostaniemy nowe równanie, tym razem na p:

(3) $$p^3 + 6p^2 + 10p - 1 = 0.$$

Skoro $x \approx 2$, to w konsekwencji $p \approx 0$. Nieznaną wartość p możemy więc wyznaczyć *w przybliżeniu*, pomijając w (3) wszystkie wyższe potęgi i dostając równanie uproszczone: $10p - 1 = 0$, skąd oczywiście $p = \frac{1}{10}$.

A dalej – wiadomo, powtórzymy schemat: teraz przecież chcemy rozwiązać równanie (3), dla którego znamy przybliżone rozwiązanie $p \approx \frac{1}{10}$. Zapisujemy więc dokładne p w postaci $p = \frac{1}{10} + q$, gdzie $q \approx 0$. Podstawiamy, dostajemy kolejny wielomian itd.

Ten algorytm dodatkowo komplikuje fakt, że na każdym kroku trzeba rozważyć nowy wielomian. To, że proces można iterować, podstawiając do *oryginalnego* wielomianu kolejno uzyskiwane przybliżenia, zauważył Joseph Raphson w 1690 roku – i dlatego metodę stycznych niektórzy nazywają *metodą Newtona–Raphsona.*

We współczesnym języku dostajemy ciąg kolejnych przybliżeń (ograniczamy się do podania wyniku do pierwszej niedokładnej cyfry):

$$\begin{aligned}
&x_0 && = 2,\\
&x_1 = x_0 + p && = 2{,}1,\\
&x_2 = x_0 + p + q && = 2{,}09456,\\
&x_3 = \ldots && = 2{,}0945514816,\\
&x_4 = \ldots && = 2{,}09455148154232659149,\\
&x_5 = \ldots \text{ itd.}
\end{aligned}$$

Jak łatwo zauważyć, nieliniowe równanie na idealną poprawkę zastępujemy równaniem liniowym, które faktycznie odpowiada znalezieniu miejsca zerowego stycznej... a więc – to *jest* metoda Newtona, zastosowana do $f(x) = x^3 - 2x - 5$. Tyle że w algorytmie podanym przez Newtona nie występuje ani wzór (1), ani nawet pojęcie pochodnej – którego przecież Newton był współtwórcą!

Z drugiej strony, musimy być świadomi przełomu, jakiego dokonał Newton: wcześniejsze metody przybliżonego rozwiązywania równań wielomianowych – na przykład metoda Viète'a z 1600 roku (której nie przedstawiamy tutaj, choć wiadomo, że Newton bardzo dokładnie ją przestudiował) – nie dość, że były bardziej złożone, to po każdej iteracji zwiększały dokładność wyniku tylko o jedną cyfrę. Tymczasem, jak mogliśmy zobaczyć w przykładzie powyżej, metoda Newtona w „typowym" przypadku po każdej iteracji z grubsza *podwaja*(!) liczbę dokładnych cyfr wyniku – i tego faktu Newton był w pełni świadom.

Ścisłą odpowiedź na pytanie o to, jak szybko jest zbieżna metoda Newtona i przy jakich założeniach na funkcję f, podał dopiero Augustin Louis Cauchy w 1829 roku. Np. dla złośliwie dobranego wielomianu $f(x) = x^2 - 2x + 1$ metoda Newtona zmniejsza błąd dwukrotnie, co mniej więcej co trzy iteracje dodaje ledwo jedną nową cyfrę wyniku.

I choć w żadnym ze znanych rękopisów Newtona nie znaleziono wzoru (1), to jednak pamiętajmy, że zdarzyło mu się użyć dokładnie takiej metody do wyznaczenia miejsca zerowego funkcji nie będącej wielomianem (zob. G. Łukaszewicz: *Równanie Keplera w Principiach Newtona*, Δ^{10}_{21}).

Metoda Newtona, zanim narodził się Newton

Czy Newton był pierwszym, który wskazał tak szybko zbieżną metodę rozwiązywania równań? Okazuje się, że w pewnych *bardzo szczególnych* przypadkach *już starożytni* korzystali z algorytmu, który dziś łatwo nam zinterpretować jako zastosowanie metody Newtona.

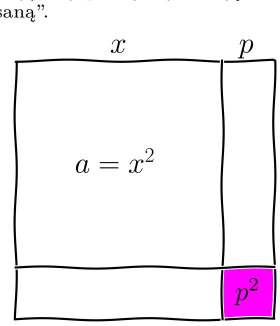
Na przykład Heron z Aleksandrii około 1600 lat przed Newtonem opisywał babilońską metodę – my znamy ją jako **metodę Herona** – wyznaczania przybliżonej wartości $x = \sqrt{a}$. Pomysł był całkiem prosty: jeśli znamy jakieś przybliżenie $x_0 > \sqrt{a}$, to wtedy $\frac{a}{x_0} < \sqrt{a}$ również jest przybliżeniem $\sqrt{a}$. Wobec tego ich średnia będzie przybliżeniem na pewno lepszym od x_0 itd... W efekcie dostajemy iterację

$$x_{k+1} = \frac{1}{2}\left(x_k + \frac{a}{x_k}\right)$$

– a to przecież nic innego, jak metoda Newtona zastosowana do równania

$$x^2 - a = 0.$$

Theon niepotrzebnie skomplikował ten proces (miał swoje powody, w które nie wnikamy); my podajemy wersję „uczesaną".



Geometryczne uzasadnienie metody Herona

Żyjący kilka wieków po Heronie Theon z Aleksandrii podawał takie oto eleganckie geometryczne uzasadnienie tego procesu. Skoro zamiast dokładnej wartości $x = \sqrt{a}$ znamy jej przybliżenie $x_0 = x + p$ (i oczywiście p jest małe), to pole kwadratu o boku x_0 jest równe

$$x_0^2 = (x + p)^2 = x^2 + 2xp + p^2 = a + 2(x_0 - p)p + p^2.$$

Zatem znów, dokładając malutki kwadracik p^2 – czyli zaniedbując wyższe potęgi p (gdzieś to chyba już widzieliśmy?) – możemy zamiast powyższego równania kwadratowego rozwiązać prostsze, liniowe:

$$x_0^2 = a + 2x_0 p,$$

skąd wyznaczamy przybliżenie idealnej poprawki $p = \dfrac{x_0^2 - a}{2x_0}$ i w konsekwencji – lepsze przybliżenie rozwiązania:

$$x_1 = x_0 - p = x_0 - \frac{x_0^2 - a}{2x_0} = \frac{1}{2}\left(x_0 + \frac{a}{x_0}\right).$$

A dalej – wiadomo, powtarzamy całą operację z nowym przybliżeniem x_1 itd.

Warto w tym momencie dodać, że Newton, a przed nim arabscy matematycy w XII wieku – zob. [Ypma, str. 539 i 541], podali analogiczne eleganckie metody przybliżania pierwiastków dowolnego stopnia n, równoważne wykorzystaniu metody stycznych do równania $x^n - a = 0$.

Metoda Newtona dzisiaj

Omawiany algorytm (1), a zwłaszcza jego wielowymiarowy wariant opracowany przez Simpsona, nie jest bynajmniej historycznym zabytkiem: to jedna z powszechnie używanych w nauce i technice metod numerycznych. Sprowadzenie trudnego zadania nieliniowego o wielu milionach niewiadomych do sekwencji zadań liniowych pozwala użyć np. wyrafinowanych algorytmów algebry liniowej (o jakich pisaliśmy w *Delcie* Δ^{7}_{18}).

Ale nie tylko! Na przykład w procesorze Intel Itanium operacja dzielenia dwóch liczb rzeczywistych b/a sprowadza się do pomnożenia $b \cdot \frac{1}{a}$. Aby zaś obliczyć $x = \frac{1}{a}$, projektanci zdecydowali się wykonać kilka iteracji algorytmu, w którym na szczęście *nie występuje* operacja dzielenia, a za to można skorzystać z instrukcji FMA (*fused multiply–add*) procesora:

$$r_k = 1 - ax_k,$$
$$x_{k+1} = x_k + x_k r_k.$$

Pozostawiamy Czytelnikowi sprawdzenie, że tak określony ciąg (x_k) jest po prostu ciągiem generowanym metodą Newtona dla $f(x) = \frac{1}{x} - a$ oraz, że dla $a > 0$ zachodzi

$$\left|x_{k+1} - \frac{1}{a}\right| = a \cdot \left|x_k - \frac{1}{a}\right|^2.$$

I właśnie na tym – że (w typowym przypadku) błąd w następnym kroku jest rzędu kwadratu błędu w bieżącym – polega nieprzemijający czar metody Newtona.

Źródła:
J.-L. Chabert, *History of Algorithms*, Springer 1999.
H. Goldstine, *A History of Numerical Analysis from the 16th through the 19th Century*, Springer 1977.
T. Ypma, *Historical Development of the Newton-Raphson Method*, SIAM Review 37 (4) 1995.

Coraz bliżej kota

Marcin BRAUN

O kocie Schrödingera słyszeli wszyscy. Wrócimy do niego za chwilę, na razie jednak przypomnijmy dwa doświadczenia, które pozwoliły odpowiedzieć na powtarzane od wieków pytanie: czy światło to fala, czy strumień cząstek. Pierwsze z nich wykonał na początku XIX wieku Thomas Young i udowodnił, że światło ulega dyfrakcji i interferencji, co świadczy niezbicie o jego falowej naturze. Kilkadziesiąt lat później inne doświadczenia pozwoliły odkryć efekt fotoelektryczny, który na początku XX wieku wyjaśnił Albert Einstein. Stwierdził on, że światło składa się z kwantów, czyli „porcji energii", zwanych dziś fotonami. W naszych rozważaniach skupimy się jednak na tym, co było później.

Fotony interferują między sobą?

Koncepcja Einsteina pozwoliła wyjaśnić efekt fotoelektryczny, a także kilka innych zjawisk. Stworzyła jednak nowy problem: jak wyjaśnić interferencję światła? Może jest ona wynikiem oddziaływania fotonów? Czy tam, gdzie powstaje ciemny prążek interferencyjny, fotony zderzają się tak nieszczęśliwie, że ulegają zniszczeniu? To można sprawdzić. Wykonajmy doświadczenie Younga jeszcze raz, ale tym razem weźmy źródło światła tak słabe, żeby wypuszczało z siebie po jednym fotonie co sekundę. Zamiast ekranu użyjemy kliszy fotograficznej, po każdym fotonie zostanie jedna kropka. Pojedynczy foton nie będzie miał z czym oddziaływać. Pewnie część fotonów przeleci przez jedną szczelinę, a część przez drugą. Nie zaobserwujemy więc prążków interferencyjnych, ale dwie plamy – po jednej za każdą szczeliną. Jeśli ekran znajduje się daleko od szczelin, plamy te nałożą się na siebie i wtedy także prążków nie będzie. Takie doświadczenie rzeczywiście zrobiono. Okazało się, że po każdym fotonie zostawała jedna kropka. W miarę jednak, jak kropek przybywało, układały się one w dobrze znane prążki interferencyjne. Tak więc foton nie interferuje z innymi fotonami. On interferuje sam ze sobą! Najwyraźniej nie tylko strumień fotonów, ale nawet pojedynczy foton jest falą.

De Broglie: wszystko jest falą

Sprawa skomplikowała się jeszcze bardziej, gdy francuski książę Louis Victor Pierre Raymond de Broglie† z książęcym iście rozmachem uznał, że skoro światło może być jednocześnie falą i strumieniem cząstek, to taką samą podwójną naturę powinny mieć wszystkie inne cząstki mikroświata. A jaką długość ma fala związana z cząstką? De Broglie skorzystał ze wzoru na pęd fotonu $p = h/\lambda$, czyli pęd $= \frac{\text{stała Plancka}}{\text{długość fali}}$. Zgodnie z tym wzorem foton o pędzie p ma długość fali $\lambda = p/h$. Uczony przyjął, że taki sam wzór obowiązuje również dla innych cząstek. W jaki sposób można sprawdzić doświadczalnie tę hipotezę? Użyjmy stosunkowo lekkich cząstek, aby przy rozsądnej prędkości miały w miarę mały pęd, a tym samym – w miarę dużą długość fali. Dobrym rozwiązaniem okazują się elektrony. Odpowiadająca im długość fali jest mniej więcej rozmiarów atomu. Co prawda trudno byłoby wyciąć dwie szczeliny oddalone o rozmiar atomu, a tym bardziej wykonać siatkę dyfrakcyjną zawierającą wiele rozmieszczonych tak szczelin, ale takie siatki Natura produkuje w wielkiej ilości. To kryształy. Już w 1927 roku, zaledwie cztery lata po pomyśle de Broglie'a, potwierdzono doświadczalnie dyfrakcję i interferencję elektronów. Później przeprowadzono to doświadczenie w nowej wersji. Dobrano prędkość elektronów i częstotliwość

†Wszyscy czytają to nazwisko /de broj/, tak też podaje francuska Wikipedia, cytując podręcznik fonetyki. Słownik Larousse'a podaje wymowę /de brogli/.

Stała Plancka wynosi:
$h = 6{,}62607015 \cdot 10^{-34}\ \mathrm{kg \cdot m^2/s}$.