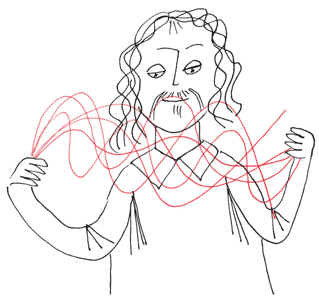


Reguła znaków Kartezjusza

Michał TARNOWSKI

Pierwiastki wielomianów są trochę jak Bóg w jednym z cytatów Einsteina: wyrafinowane, ale nie złośliwe. Z jednej strony w ogólności nie ma na nie „jawnych wzorów”, ale z drugiej wiadomo, że te pierwiastki zawsze istnieją, przynajmniej jeśli rozszerzymy poszukiwania do liczb zespolonych.

W tym artykule ograniczymy poszukiwania do pierwiastków rzeczywistych, co bardziej odpowiada szkolnym realiom. Czasem nieskomplikowane reguły pozwalają czegoś się o nich dowiedzieć. Najprostsze z tych kryteriów są tak intuicyjne, że nie mają swojej nazwy. Przedstawimy teraz dwa z nich.

Wykluczanie półprostej

Jeśli wszystkie niezerowe współczynniki wielomianu są tego samego znaku (np. x^5+2x+1 lub $-x^6-x^2-3$), to próżno oczekiwać rozwiązań dodatnich, bo dodatnie argumenty prowadzą do wartości o znaku takim, jaki mają współczynniki wielomianu. Można powiedzieć, że daje to nam warunek wystarczający braku pierwiastków dodatnich. Nasuwa się pytanie: czy jest on również konieczny? Odpowiedź brzmi: nie; kontrprzykładem jest tu choćby $x^2-2x+2=(x-1)^2+1$, pozbawiony jakichkolwiek pierwiastków rzeczywistych. Jednocześnie, jeśli wielomian ma tyle pierwiastków ujemnych, ile wynosi jego stopień, to jego współczynniki są jednakowego znaku, co wynika wprost ze wzorów Viète'a.

Podobne rozumowanie pozwala czasem wykluczyć pierwiastki ujemne, bo są to dodatnie pierwiastki lustrzanego odbicia wielomianu względem osi pionowej: $f(-x)$. Przykładowo funkcja $f(x)=x^8-x+1$ nie ma ujemnych miejsc zerowych, bo dla ujemnych argumentów przyjmuje dodatnie wartości.

Gwarancje istnienia na półprostej

Dla wielomianu $f(x)=a_nx^n+\dots+a_1x+a_0$ wyraz wolny a_0 to inaczej jego wartość w zerze, za to współczynnik wiodący a_n mówi coś o granicy tej funkcji w nieskończoności: $\operatorname{sgn} a_n=\operatorname{sgn}\lim_{x\to\infty} f(x)$. Dlatego jeśli znaki skrajnych współczynników są przeciwne ($a_0a_n<0$), to wielomian gdzieś na półprostej dodatniej musi mieć miejsce zerowe – zgodnie z twierdzeniem Darboux o przyjmowaniu wszystkich wartości pośrednich przez funkcję ciągłą.

Z przedstawionych wyżej kryteriów można wywnioskować na przykład, że wielomian x^7+x-2 nie ma pierwiastków ujemnych, za to musi mieć co najmniej jeden dodatni.

Czy to koniec listy heurystyk? Czy znaki współczynników mogą mieć jeszcze jakiś inny układ, który wyklucza pierwiastki z którejś półprostej lub przeciwnie – gwarantuje jakieś pierwiastki na niej? Tu na scenę wkracza tytułowa *reguła znaków* przypisywana Kartezjuszowi – mimo że pierwszy dowód jej poprawności może pochodzić dopiero od późniejszych matematyków. Reguła znaków Kartezjusza (RZK) wiąże ze sobą dwie proste wielkości:

- p, czyli liczbę dodatnich pierwiastków niezerowego wielomianu, licząc ich krotności,
- s, czyli liczbę zmian znaków przy jego niezerowych członach, uporządkowanych „standardowo”, według malejących potęg.

RZK orzeka, że **różnica $s-p$ jest nieujemna i parzysta.**

Widać, że wynikają z niej wcześniej omawiane kryteria:

- jednakowe znaki przy wszystkich niezerowych członach, czyli $s=0$, dają $p\leqslant 0$, czyli $p=0$ – brak pierwiastków dodatnich;
- jeśli współczynnik wiodący i wyraz wolny mają przeciwne znaki ($a_0a_n<0$), to liczba s jest nieparzysta, a zatem p również musi takie być i wynosi co najmniej jeden.

Przypomnijmy wzory Viète'a: jeśli $x_1,\dots,x_n$ są pierwiastkami wielomianu $P(x)=\sum_{i=0}^n a_ix^i$, to zachodzi:

$$\frac{a_{n-1}}{a_n}=-\sum_{i\leqslant n}x_i,$$
$$\frac{a_{n-2}}{a_n}=\sum_{i<j\leqslant n}x_ix_j,$$
$$\dots,$$
$$\frac{a_0}{a_n}=(-1)^n\prod_{i\leqslant n}x_i.$$

Rozwiązanie zadania M 1747.
Niech $p_1,p_2,\dots,p_{2023}$ będą różnymi liczbami pierwszymi. Wówczas zbiór $\{a_1,a_2,\dots,a_{2023}\}$, gdzie

$$a_i=\frac{1}{p_i}\cdot p_1p_2\dots p_{2023}$$

spełnia warunki zadania.

Rozwiązanie zadania M 1748.
Zauważmy, że gdyby przy stole siedzieli sami kłamcy, nikt z nich nie mógłby powiedzieć: „Wszyscy, których widzę, są kłamcami”, ponieważ byłaby to prawda. Zatem przy stole siedzi przynajmniej jeden rycerz, nazwijmy go Ahmed. Skoro jest on prawdomówny, to wszystkich 27 osób, które widzi, jest kłamcami.

Gdyby jego obaj sąsiedzi byli kłamcami, to mówiliby prawdę, deklarując, że widzą samych kłamców – co jest niemożliwe. Zatem jeden z sąsiadów Ahmeda również jest rycerzem. Każda inna osoba siedząca przy stole poza tymi dwoma rycerzami jest widziana przez Ahmeda lub jego sąsiada rycerza, więc musi być kłamcą. Nietrudno się również przekonać, że konfiguracja, w której występuje tylko dwóch siedzących obok siebie rycerzy, faktycznie spełnia warunki zadania.

Pozostaje przejść do dowodu ogólnego twierdzenia; mówiąc matematycznym żargonem: jest on elementarny, ale nietrywialny.

Rozpoczniemy od uzasadnienia, że $s - p$ jest liczbą parzystą. Możliwe są dwa przypadki: wyrazy wiodący i wolny mają albo takie same znaki ($a_n a_0 > 0$), albo przeciwne ($a_n a_0 < 0$). W pierwszym wypadku s i p są oba parzyste (druga parzystość wynika z tego, że wartości wielomianu w 0 i w granicy ∞ mają ten sam znak). Analogicznie, w drugim przypadku obie wielkości muszą być nieparzyste. Obie ewentualności dają parzystą wartość różnicy $s - p$.

Pozostaje udowodnić, że $s - p \geqslant 0$. Ta część dowodu korzysta z kilku elementów: własności pochodnych wielomianów, z indukcji, z twierdzenia Rolle'a oraz z części pierwszej. Porównanie wielkości s i p jest pośrednie, przez analogiczne parametry wielomianu pochodnego: s', p'.

(i) Różniczkowanie wielomianu nie zmienia znaku współczynników, licząc od tego przy najwyższej potędze: $(a_k x^k)' = k a_k x^{k-1}, k \geqslant 0$. W ten sposób na pewno nie powstanie żadna nowa zmiana znaku. Może się co najwyżej zdarzyć, że jedna ze zmian znaku wiązała się z wyrazem wolnym, który znika przy różniczkowaniu; zatem $s \geqslant s'$.

(ii) Dowodzona nierówność $s \geqslant p$ w trywialny sposób zachodzi dla wielomianów stopnia 0 (tzn. stałych), wtedy $s = p = 0$. Dlatego aby udowodnić, że nierówność zachodzi w ogólności, wystarczy pokazać krok indukcyjny – czyli że nierówność dla ustalonego stopnia n wynika z nierówności dla niższych stopni. Ponieważ różniczkowanie zmniejsza stopień wielomianu o 1, z założenia indukcyjnego mamy $s' \geqslant p'$.

(iii) Zgodnie z twierdzeniem Rolle'a (rys. 1) pomiędzy pierwiastkami wielomianu muszą się znajdować pierwiastki jego pochodnej. Odpowiednich przedziałów jest o jeden mniej niż pierwiastków f, którymi te przedziały są rozdzielone i otoczone; w takim razie $p' \geqslant p - 1$.

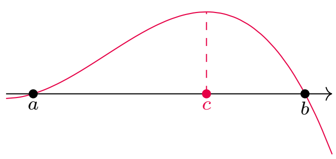
Te trzy spostrzeżenia pozwalają ułożyć ciąg nierówności: $s \geqslant s' \geqslant p' \geqslant p - 1$. Prowadzi to do wniosku, że $s - p \geqslant -1$, a udowodniona wcześniej parzystość tej różnicy pozwala poprawić tę nierówność: $s - p \geqslant 0$. Krok indukcyjny został wykazany, a razem z nim ogólne twierdzenie. $\square$

Oczywiście reguła Kartezjusza pozwala też powiedzieć coś o liczbie pierwiastków ujemnych – wystarczy rozważać wielomian z odwróconym znakiem argumentu, $f(-x)$. Dalsze zmiany współrzędnych pozwalają też powiedzieć coś o liczbie miejsc zerowych w dowolnie ustalonym przedziale – takie uogólnienie bywa nazywane twierdzeniem Budana lub Fouriera.

Dla przykładu rozważmy wielomian $p(x) = x^3 - 5x + 1$. Zgodnie z regułą Kartezjusza ma on 0 lub 2 dodatnie pierwiastki. Przyjrzyjmy się wielomianowi

$$q(y) = (y+1)^3 p\left(\frac{3y+2}{y+1}\right) = 13y^3 + 17y^2 + 4y - 1.$$

Funkcja $\varphi(y) = \frac{3y+2}{y+1} = 3 - \frac{1}{y+1}$ jest rosnąca na półprostej dodatniej, z czego wynika, że każdemu dodatniemu pierwiastkowi y^* wielomianu $q(y)$ odpowiada jednoznacznie pierwiastek $x^* = \varphi(y^*)$ wielomianu $p(x)$, należący do przedziału $\big(\varphi(0), \varphi(\infty)\big) = (2, 3)$. Zgodnie z RZK wielomian $q(y)$ ma dokładnie jeden dodatni pierwiastek, zatem wielomian $p(x)$ ma dokładnie jeden pierwiastek w przedziale $(2, 3)$. Można powiedzieć, że pierwiastek ten został *wyizolowany* i, gdybyśmy chcieli, moglibyśmy go dalej numerycznie przybliżyć z dowolną dokładnością, korzystając na przykład z metody bisekcji (rys. 2). Tego typu analizę można przeprowadzić w sposób systematyczny, tak aby wyizolować wszystkie rzeczywiste pierwiastki zadanego wielomianu, tzn. wyznaczyć rozłączne odcinki, z których każdy zawiera dokładnie jeden pierwiastek, i jednocześnie mieć pewność, że poza wyznaczonymi odcinkami tych rzeczywistych pierwiastków nie ma. Na takim podejściu opierają się na przykład metody VAG (Vincenta–Alesiny–Galuzziego, 2000) i VAS (Vincenta–Akritasa–Strzebońskiego, 2005). Reguła znaków Kartezjusza jest zatem czymś więcej niż tylko historyczną ciekawostką.



Rys. 1. Ilustracja do twierdzenia Rolle'a. Między miejscami zerowymi funkcji różniczkowalnej (punkty a i b) znajduje się miejsce zerowe jej pochodnej (punkt c). Można je scharakteryzować jako ekstremum lokalne, którego istnienie gwarantuje twierdzenie Weierstrassa.

Bibliografia

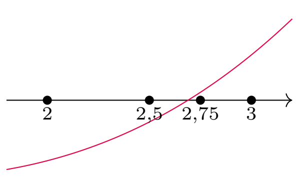
Przemysław Koprowski, *Matematyka obliczeniowa 5: izolacja pierwiastków*, YouTube, dostęp: 2020-12-01.

Przemysław Koprowski, *Lectures on Computational Mathematics*, s. 277, `http://www.pkoprowski.eu/lcm/lcm.pdf`, dostęp: 2022-05-21.

William L. Hosch, *Descartes's rule of signs*, Britannica Online.

Descartes' Rule of Signs, ProofWiki.org.

Descartes Rule of Signs, nothing but math proofs, YouTube, dostęp: 2021-01-27.



Rys. 2. Przykład działania metody bisekcji dla ciągłej funkcji $f(x)$ przyjmującej wartości różnych znaków na krańcach przedziału $(2, 3)$. Jako pierwszy wyraz iteracji bierzemy środek przedziału. Ponieważ $f(2{,}5) < 0$, zgodnie z własnością Darboux funkcji ciągłych miejsce zerowe istnieje w przedziale (2,5, 3) – i jako drugi wyraz bierzemy środek tego przedziału. Procedurę powtarzamy, mając gwarancję, że n-ty wyraz tak uzyskanego ciągu przybliżeń jest odległy od miejsca zerowego o nie więcej niż 2^{-n}