

Najlepiej dopasowana prosta?

*Instytut Fizyki PAN

Marek W. GUTOWSKI*

Tytułowa fraza – oczywiście bez znaku zapytania – jest często spotykana w publikacjach naukowych. *Wszyscy wiedzą*, co ona oznacza: chodzi o taką prostą, która jest najbliższa do jednocześnie prezentowanego zbioru wyników pomiarów. *Wszyscy wiedzą*, że parametry tejże prostej można otrzymać za pomocą metody najmniejszych kwadratów, zaproponowanej przez Gaussa. Niniejszy esej ma na celu pokazanie użyteczności *analizy interwałowej (przedziałowej)* w procesie znanym w slangu laboratoryjnym jako *fitowanie danych* (albo, jak kto woli, *fitowanie krzywych*). Jednocześnie prezentujemy nieznaną dotąd związek analizy przedziałowej z rachunkiem prawdopodobieństwa.

Identyczne rachunki – ale przeprowadzane na różnych maszynach – dają różne wyniki. Dobrze byłoby więc przynajmniej wiedzieć *na pewno*, w jakich granicach mieści się wynik prawdziwy. Jasne, że chcielibyśmy też, aby oszacowany przedział wartości był możliwie wąski. To właśnie było głównym motywem pierwszych prac w dziedzinie rachunków interwałowych. Za ojca tej części matematyki powszechnie uważa się Amerykanina, Ramona E. Moore'a (1929–2015), który ustalił jej solidne podwaliny (raport techniczny dla firmy Lockheed (styczeń 1959), rozprawa doktorska (1962), monografia *Interval Analysis* (1966)). Wprawdzie nasz rodak, Mieczysław Warmus (1918–2007), wcześniej opublikował dwa artykuły o tejże tematyce: *Calculus of Approximations* (1956) oraz *Approximations and Inequalities in the Calculus of Approximations. Classification of Approximate Numbers* (1961), ale te ukazały się w Biuletynie Polskiej Akademii Nauk, czyli za „żelazną kurtyną”, przez co nie zostały zauważone przez resztę świata. Można też zażartować, że prekurem w tej dziedzinie był Archimedes, który jako pierwszy podał obustronne oszacowanie liczby π : $3\frac{10}{71} < \pi < 3\frac{10}{70}$. Więcej o historii analizy interwałowej na stronie <http://www.cs.utep.edu/interval-comp/early.html>.

Dziś rachunki interwałowe znalazły praktyczne zastosowanie w dziedzinach, w których projektuje się bardzo kosztowne konstrukcje (np. morskie wieże wiertnicze), wymagany jest wysoki stopień bezpieczeństwa ludzi (załogowy sprzęt kosmiczny) lub gdy wymagana jest wysoka precyzja obliczeń (np. bezkolizyjne sterowanie ruchem robotów podwodnych).

Przejdźmy teraz do krótkiego opisu interwałów i operacji, które możemy na nich wykonywać. Interwał to obustronnie domknięty przedział osi liczbowej, zapisywany zwykle jako $\mathbf{x} = [\underline{x}, \bar{x}]$. Litery pisane tłustym drukiem będą odtąd oznaczać interwały, których końce będziemy oznaczać tą samą literą: dolny z podkreśleniem, a górny z kreską u góry. Zbiór wszystkich interwałów zwykle oznacza się jako \mathbb{IR} , przez analogię do oznaczenia zbioru liczb rzeczywistych jako \mathbb{R} . Mamy więc: $\mathbb{IR} = \{\mathbf{x} = [\underline{x}, \bar{x}] : \underline{x}, \bar{x} \in \mathbb{R}, \underline{x} \leq \bar{x}\}$. Zauważmy, że poprawnymi interwałami są również odcinki o zerowej długości (czyli takie, że $\underline{x} = \bar{x}$), które nazywamy *singletonami*. Możemy to zapisać krótko

jako $\mathbb{R} \subset \mathbb{IR}$ – liczby rzeczywiste są podzbiorem zbioru interwałów.

Chcemy posługiwać się interwałami w sytuacjach, kiedy nasze liczby są niepewne, ale wiemy, w jakich granicach mieszczą się one z całą pewnością. Zaczniemy od czterech działań arytmetycznych. Nietrudno się przekonać, że $\mathbf{x} + \mathbf{y} = [\underline{x} + \underline{y}, \bar{x} + \bar{y}]$, natomiast $\mathbf{x} - \mathbf{y} = [\underline{x} - \bar{y}, \bar{x} - \underline{y}]$. Mnożenie i dzielenie są bardziej pracochłonne. Niech \diamond oznacza znak mnożenia albo dzielenia. Najpierw znajdujemy czteroelementowy zbiór $\mathcal{Z} = \{\underline{x} \diamond \underline{y}, \underline{x} \diamond \bar{y}, \bar{x} \diamond \underline{y}, \bar{x} \diamond \bar{y}\}$. Wynik mnożenia, bądź dzielenia, wyrazi się wówczas jako: $\mathbf{x} \diamond \mathbf{y} = [\min \mathcal{Z}, \max \mathcal{Z}]$, przy czym przy dzieleniu wymagamy, aby $0 \notin \mathbf{y}$ (nie dzielimy przez zero).

Ważne: wyniki prostych działań arytmetycznych to zawsze interwały *ciasne*, czyli takie, które zawierają wszystkie możliwe wyniki operacji przeprowadzonych na dwóch liczbach, z których pierwsza pochodzi z interwału \mathbf{x} , a druga z interwału \mathbf{y} , i *tylko* te wyniki, bez jakichkolwiek przeszacowań. Problem jednak w tym, że nasze obliczenia będą realizowane przez komputery, które pracują ze skończoną dokładnością. W praktyce jest więc wręcz konieczne, aby wyniki nawet takich prostych obliczeń były odpowiednio zaokrąglane, mianowicie „na zewnątrz”, czyli lewy (dolny) kraniec interwału wynikowego w dół, a górny – w górę. Tylko w ten sposób nasze wyniki będą gwarantowane, choć być może przeszacowane.

Niestety, wyrażenia bardziej skomplikowane często produkują interwały szersze, choć *zawsze* zawierające te ciasne. Warto samemu przekonać się, że generalnie mamy $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) \subseteq \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$, a nie równość! Wniosek praktyczny jest następujący: zanim obliczymy wartość skomplikowanego wyrażenia, postarajmy się najpierw przepisać je do takiej postaci, aby każda zmienna pojawiała się w nim *tylko jeden raz*. Na przykład: opór zastępczy dwóch oporników, R_1 i R_2 , połączonych równoległe, zwykle wyliczamy ze wzoru $R = R_1 \cdot R_2 / (R_1 + R_2)$. W rachunkach interwałowych lepszy, bo ciasny, wynik otrzymamy, stosując równoważny wzór: $\mathbf{R} = 1 / (\frac{1}{\mathbf{R}_1} + \frac{1}{\mathbf{R}_2})$.

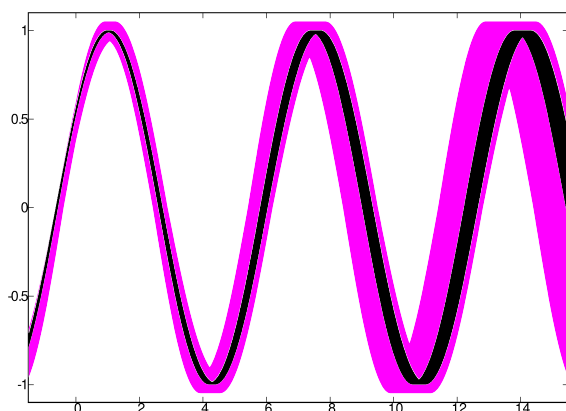
Ale przecież interwał to zbiór – powinniśmy więc mieć możliwość wykonywania operacji typowych dla zbiorów.

Nietrudno się przekonać, że $\mathbf{p} \cap \mathbf{q}$ (część wspólna) to przedział $[\max(\underline{p}, \underline{q}), \min(\overline{p}, \overline{q})]$, natomiast $\mathbf{p} \cup \mathbf{q}$ (suma) wyrazi się jako $[\min(\underline{p}, \underline{q}), \max(\overline{p}, \overline{q})]$. Tu pora na dwie uwagi:

1) Część wspólna może przecież być zbiorem pustym. Nasz wzór wyprodukuje wtedy „nielegalny” interwał, tzn. taki, że jego dolny kraniec przewyższa kraniec górny. W praktycznych rachunkach interwałowych wygodnie jest zapisać zbiór pusty (interwał pusty) w postaci $\emptyset = [+INF, -INF]$, gdzie INF jest największą liczbą maszynową. Wprawdzie klóci się to z wymaganiami, aby spełniona była nierówność $\underline{x} \leq \overline{x}$, ale bardzo praktyczne jest zastąpienie każdego wykrytego „nielegalnego” wyniku tak właśnie zapisanym interwałem pustym.

2) Teoriomnogościowa suma dwóch interwałów niekoniecznie jest pojedynczym interwałem. Tymczasem w takim wypadku nasz wzór dostarczy nam obiekt będący interwałem, zawierającym oprócz wszystkich liczb zawartych w \mathbf{p} lub \mathbf{q} także inne liczby. Nazywamy taki obiekt *powłoką wypukłą* interwałów \mathbf{p} i \mathbf{q} .

Pora na coś więcej. Jak obliczyć wartość pewnej funkcji f , której argumentem jest interwał \mathbf{x} ? Powinien to być interwał będący zbiorem wszystkich wartości, jakie przybiera funkcja f , gdy jej argument zmienia się w przedziale \mathbf{x} . Łatwo znaleźć dokładny wynik, jeśli na odcinku \mathbf{x} funkcja f jest stała, rosnąca lub malejąca. W innych przypadkach sprawy się komplikują (wynik może być przeszacowany), jednakże dostępne oprogramowanie radzi sobie z tym. Zawsze jednak z warunku $\mathbf{x}_1 \subset \mathbf{x}_2$ wynika, że $f(\mathbf{x}_1) \subseteq f(\mathbf{x}_2)$. Własność tę ilustruje rysunek 1.

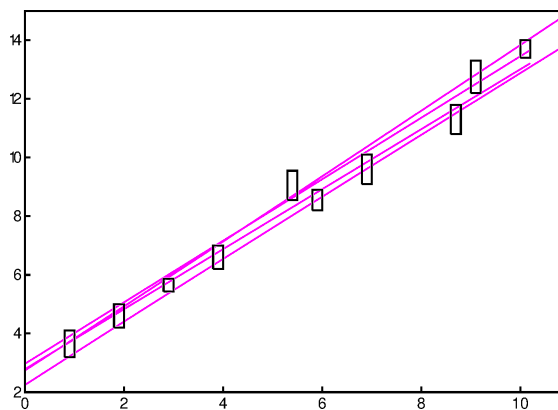


Rys. 1. Wykres funkcji $f(t) = A \cdot \sin(\omega t + \varphi)$. Czas t jest zwykłą liczbą rzeczywistą, ale pozostałe parametry są niepewne: amplituda $A = [0.95, 1.05]$, częstość $\omega = [0.95, 1.05]$, faza $\varphi = [0.40, 0.60]$ (krzywa kolorowa). Po zwężeniu przedziałów parametrów do $A = [0.99, 1.00]$, $\omega = [0.95, 0.98]$ i $\varphi = [0.55, 0.60]$ otrzymujemy czarną krzywą, która – oczywiście – w całości zawiera się w kolorowej krzywej

Przemyciliśmy na nim coś, co można by nazwać interwałami wielowymiarowym. W literaturze takie obiekty zwykle nazywane są *kostkami* lub *pudełkami*. Czym bowiem jest trójwymiarowy interwał (A, ω, φ) z rysunku 1? Jeśli zechcemy go zilustrować w prostokątnym układzie współrzędnych, to zobaczymy prostopadłościan, którego krawędzie będą równoległe do osi układu – co uzasadnia tę nazwę, używaną także

przy większej lub mniejszej liczbie wymiarów. Przy okazji zdefiniujemy jeszcze pojęcie *średnicy* kostki. Jest to po prostu długość jej najdłuższej krawędzi. Na tym kończymy prezentację narzędzi potrzebnych w tym eseju. Więcej m.in. w Wikipedii, pod hasłem `Interval_arithmetic`.

Opiszemy teraz dwa algorytmy rzetelnego szacowania wartości liczbowych parametrów fizycznych oraz ich niepewności, ukrytych w pracowni zebranych danych. Rysunek 1 sugeruje niedwuznacznie, że powinniśmy wystartować ze względnie dużej kostki, co do której mamy pewność, że zawiera prawdziwe wartości szukanych parametrów. To może być problem sam w sobie, ale w przypadku pasowania linii prostej, takiej jak na rysunku 2, sprawa nie wydaje się trudna.



Rys. 2. Przykładowe dane eksperymentalne z niepewnościami w obydwu współrzędnych oraz kilka linii prostych przechodzących przez możliwie wiele „prostokątów niepewności”. Która z nich jest najlepsza?

Widzimy, że jeśli nasza prosta ma równanie $y = ax + b$, to $\mathbf{a} = [0.5, 2]$ oraz $\mathbf{b} = [-2, 6]$ będą dobrymi wstępnymi oszacowaniami poszukiwanych parametrów a i b . My dla wszelkiej pewności, wystartowaliśmy z wyraźnie większej kostki: $a \in [-2.375, 12.5]$ oraz $b \in [-32.45, 34.3]$.

Pierwszy algorytm polega na systematycznym zmniejszaniu startowej kostki poprzez odcinanie od niej „plasterków”. Najpierw odcinamy połowę aktualnej kostki, na przykład dolną połowę przedziału parametru a . Potem testujemy odcięty plasterzek poprzez sprawdzenie, czy generowane z niego linie proste mają części wspólne z którymkolwiek wynikiem pomiarów. Jeśli tak, to cięcie się nie udało i wobec tego próbujemy odciąć cieńszy plasterzek, mianowicie o połowę cieńszy. Jeśli cięcie okazało się skuteczne, to w dalszych etapach będziemy pracować już ze zmniejszoną kostką. I tak aż do skutku, czyli:

a) aż odcinana kostka nie będzie generować ani jednej prostej choćby tylko zawadzającej o oryginalne dane, albo

b) wykonamy bezskutecznie maksymalną dopuszczalną liczbę prób. To znaczy ile? Hmm, to zależy od nas i od dokładności maszynowej. Nie ma sensu wykonywanie większej liczby cięć niż liczba bitów słowa maszynowego. Warto jednakże zdać sobie sprawę z tego, że już dziesiąta próba to odcięcie $\simeq 1/1000$ szerokości

aktualnie testowanego przedziału, czyli zmiana 3-4 cyfry znaczącej.

Te czynności powtarzamy dla górnej granicy parametru a , a potem także dla parametru b . Opisany cykl powtarzamy do momentu, kiedy nie uda się poprawić żadnego krańca wszystkich poszukiwanych parametrów. Wynikiem jest „zaledwie” powłoka interwałowa poszukiwanych parametrów. Dla prostej z rysunku 2 otrzymaliśmy $\mathbf{a} = [1.02271, 1.04840]$, $\mathbf{b} = [2.78378, 2.96827]$ (po zaokrągleniu do 5 cyfr znaczących). Warto porównać ten wynik z otrzymanym metodą najmniejszych kwadratów, uwzględniającą niepewności w obydwu współrzędnych: $a = 1.08663392 \pm \pm 0.0136490939$, $b = 2.49181199 \pm 0.082284525$ (spisane prosto z ekranu, bez żadnych zaokrągleń). Ten drugi wynik jest podany w dobrze znanym formacie: *wartość średnia \pm odchylenie standardowe*. Obydwa wyniki są podobne, choć nasz interwałowy jest jakby nieco szerszy. Warto zauważyć, że tzw. najlepiej dopasowana prosta ma parametry *spoza* kostki interwałowej. Nie jest to pierwszy przypadek tego rodzaju opisany w literaturze.

Dygresja: Ten sposób dopasowania przyda się głównie podczas kalibracji i do określenia dokładności przyrządu pomiarowego, kiedy jego niepewności są zupełnie nieznanne i mamy do dyspozycji jedynie „surowe” wyniki pomiarów. Wówczas najlepiej dopasowana prosta to taka, że odległość najbardziej odstającego od niej punktu jest najmniejsza z możliwych. Ale to *nie* jest taka sama prosta, jaką otrzymalibyśmy z metody najmniejszych kwadratów! Nie wolelibyśmy przypadkiem najbardziej prawdopodobnej (wiarygodnej) prostej?

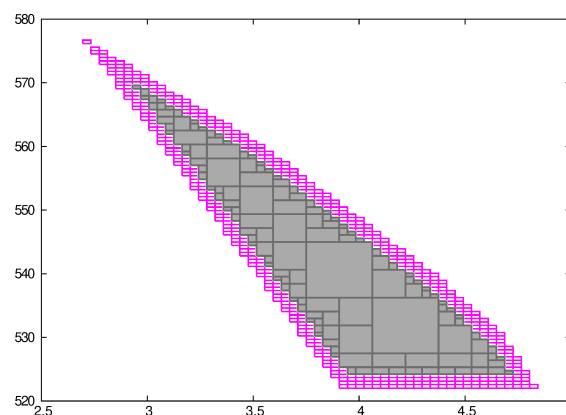
Drugi algorytm operuje na liście kostek, która na początku zawiera tylko jeden element – kostkę startową. Musimy jednak, oprócz zwykłych danych wejściowych, poinformować dodatkowo komputer, jakiej dokładności oczekujemy dla każdego z szukanych parametrów. Będą to jednostki długości potrzebne do określania średnic kostek.

Procedura jest następująca: wybieramy największą kostkę z listy (ona znika z listy) i dzielimy ją na dwie połówki. Podział odbywa się oczywiście na najdłuższej krawędzi. Teraz badamy obydwie połówki. Badana kostka jest „dobra”, jeśli generuje „grube” linie proste mające potencjalne części wspólne z więcej niż połową danych eksperymentalnych. Jeśli tak, to dopisujemy ją do naszej listy, w przeciwnym wypadku zapominamy ją. Proces się kończy, kiedy albo lista okaże się pusta, albo będą na niej tylko małe kostki, czyli mające średnicę co najwyżej równą 1.

Pusta lista to zapewne wynik niewłaściwego wyboru początkowego obszaru poszukiwań. Możliwe jest jednak, że wybrany model (w tym przypadku linia prosta) po prostu nie nadaje się do opisu badanego zjawiska. Zwykle jednak wygenerujemy grono kostek podobne do tego z rysunku 3. Teraz już nietrudno wyliczyć średnie wartości poszukiwanych parametrów oraz ich odchylenia standardowe. W tym celu wystarczy potraktować każdy punkt przynależny do znalezionej grono jako wynik

prostego, bezpośredniego pomiaru. Zainteresowani wartościami współczynników korelacji też sobie poradzą.

Pozostaje wyjaśnić, dlaczego za dobre uznajemy kostki generujące linie zaczepiające o więcej niż połowę zebranych punktów – a nie o wszystkie punkty, oraz znaczenie koloru szarego na rysunku 3. Otóż kolor szary oznacza kostki generujące linie proste na tyle „chude”, że *w całości* mieszczą się w więcej niż połowie eksperymentalnych bramek niepewności. Takich kostek nie warto dzielić na mniejsze części, co może zaoszczędzić sporo czasu i pamięci komputera.



Rys. 3. Surowy wynik interwałowego dopasowania linii prostej do danych eksperymentalnych (innych niż te na rys. 2). Dowolny punkt z obszaru szarego opisuje prostą, która *na pewno* przecina więcej niż połowę „prostokątów niepewności”. Punkty kolorowe nie mają tej gwarancji, są kategorii „być może”

A dlaczego nie wymagamy, aby dopasowana prosta przechodziła przez wszystkie punkty pomiarowe? Otóż wynik pomiaru w formie $y \pm \sigma$ wcale nie oznacza, że wartość prawdziwa mieści się w przedziale $[y - \sigma, y + \sigma]$. Przecież *wszyscy wiedzą*, że zdarzają się *pomiary odstające* (błędny odczyt/zapis, przekłamanie transmisji, jakieś zakłócenia pomiaru itp.), dawniej znane pod nazwą *błędów grubych*. Jeśli rozkład niepewności pomiarowych jest normalny (gaussowski), to wartość prawdziwa ma około 67% szans na znalezienie się w tym przedziale. Przy nieznanym rozkładzie niepewności pomiarowych musimy się podeprzeć twierdzeniem Czebyszewa (1874), które mówi, że prawdopodobieństwo odchylenia od wartości średniej o więcej niż $k\sigma$ ($k > 1$) nie przekracza $1/k^2$. Inaczej mówiąc, w przedziale $[y - k\sigma, y + k\sigma]$ powinno się mieścić $100 \cdot (1 - 1/k^2)$ procent pomiarów. W naszych dopasowaniach zamiast przedziałów $[y - \sigma, y + \sigma]$ powinniśmy więc używać innych (szerszych): $[y - k\sigma, y + k\sigma]$, ze współczynnikiem rozszerzenia k tak dobranym, aby prawdopodobieństwo przejścia dopasowanej prostej przez więcej niż połowę zebranych pomiarów przekraczało 1/2. Wydawałoby się, że $k = \sqrt{2}$ powinno być dobre w każdym przypadku. Tak jest, o ile liczba pomiarów N jest bardzo duża. Dla $N > 55$ dobre są oszacowania (dolne): $k > \sqrt{2}(1 + \frac{3}{2N})$ (dla N parzystych) oraz $k > \sqrt{2}(1 + \frac{1}{N})$ (dla N nieparzystych). Tymczasem dla $N = 3, 6, 9, 30$ mamy odpowiednio: $k = 2.201664, 1.944591, 1.594986, 1.489787$. Na zakończenie: praktyka pokazuje, że prosta dopasowana ostatnią metodą zwykle przechodzi przez 70–90% prostokątów niepewności punktów pomiarowych.