

Przez trzy kolejne dni mierzyłem odległość od mojego domu do najbliższego kiosku z gazetami i otrzymałem następujące wyniki (w krokach): 348, 352 i 342. Pytanie: jaka jest odległość od mojego domu do tego kiosku?

A oto mniej błahy problem o dokładnie takiej samej strukturze. Dla ustalenia ładunku elektronu dokonano pięciu niezależnych pomiarów i otrzymano wyniki (w absolutnych jednostkach elektrostarycznych): $4,781 \cdot 10^{-10}$; $4,795 \cdot 10^{-10}$; $4,769 \cdot 10^{-10}$; $4,792 \cdot 10^{-10}$ i $4,779 \cdot 10^{-10}$. Pytanie: ile wynosi ładunek elektronu? Co w ogóle można na ten temat powiedzieć mając takie wyniki pomiarów?

Wróćmy do pomiarów, które każdy z nas może wykonać samodzielnie i rozpatrzmy trochę prostszą wersję naszego pierwszego zadania. Przypuśćmy, że zmierzaliśmy odległość od domu do najbliższego kina i otrzymaliśmy wynik 812 kroków. Co na tej podstawie możemy powiedzieć o „prawdziwej” odległości do kina, skoro wiemy, że drugi pomiar dałby już inny wynik? Oto pewne, może nieco zbyt egocentryczne, rozumowanie prowadzące do pewnej odpowiedzi. W wyniku pomiaru interesującej nas wielkości, powiedzmy a , otrzymaliśmy wynik, powiedzmy x . Popelniliśmy błąd $\varepsilon = x - a$, ale oczywiście nie znamy wielkości tego błędu. Chcieliśmy wykonać nasz pomiar jak najstaranniej, to znaczy chcieliśmy wykonać go tak, żeby błąd był jak najbardziej bliski zeru. Przypuśćmy, że udało się nam to osiągnąć. W konsekwencji tego założenia za a powinniśmy przyjąć taką liczbę, która przy ustalonym wyniku pomiaru x minimalizuje $|x - a|$. Taką liczbą jest oczywiście x , a więc w naszym przykładzie stwierdzamy, że odległość od domu do najbliższego kina wynosi 812 kroków. Zasadę, na której oparliśmy oszacowanie interesującej nas wielkości, mogliśmy z oczywistych powodów nazwać zasadą najmniejszego błędu. Sprawa nieco komplikuje się, gdy — jak w pierwszym naszym zadaniu — wykonujemy więcej niż jeden pomiar. Jeżeli przez a oznaczymy odległość od domu do kiosku oraz przez $\varepsilon_1, \varepsilon_2, \varepsilon_3$ — błędy jakie popelniliśmy w kolejnych pomiarach, to mamy

$$a + \varepsilon_1 = 348$$

$$a + \varepsilon_2 = 352$$

$$a + \varepsilon_3 = 342$$

Jak teraz oszacować wielkość a ? Postąpimy znowu zgodnie z zasadą najmniejszego błędu. Zauważmy, że teraz — w trzech pomiarach — popelniliśmy trzy błędy: $\varepsilon_1, \varepsilon_2, \varepsilon_3$, wygodniej jednak będzie mówić o błędzie $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ i o jego „składowych” $\varepsilon_1, \varepsilon_2, \varepsilon_3$. Ta terminologia sugeruje, żeby w tym przypadku błąd reprezentować za pomocą punktu w przestrzeni trójwymiarowej i żeby za wielkość błędu przyjąć odległość tego punktu od zera (od początku układu współrzędnych). Jeżeli umówimy się, że odległość w tej przestrzeni trójwymiarowej będzie zwykłą odległością euklidesową, to zasada najmniejszego błędu nakazuje nam ustalić a tak, żeby $\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2$ było możliwie najmniejsze, to znaczy tak, żeby zminimalizować wielkość:

$$(a - 348)^2 + (a - 352)^2 + (a - 342)^2.$$

Taka metoda szacowania nieznannej wielkości a nazywa się — znowu z oczywistych powodów — metodą najmniejszych kwadratów. Czytelnik zechce sprawdzić, że metoda ta daje wynik $a = (348 + 352 + 342)/3$.

Uogólnienie na n pomiarów jest proste: jeżeli x_1, x_2, \dots, x_n są wynikami pomiaru pewnej wielkości a , to za oszacowanie tej wielkości metodą najmniejszych kwadratów przyjmuje się po prostu średnią arytmetyczną $(x_1 + x_2 + \dots + x_n)/n$. Jest to zgodne z oczekiwaniem. A oto mniej banalne przykłady zastosowania metody najmniejszych kwadratów.

Przykład 1. Ze szkolnego kursu fizyki wszyscy znamy prawo Boyle'a-Mariotte'a: iloczyn objętości i ciśnienia gazu ma w tej samej temperaturze wartość stałą. Zapisuje się to za pomocą znanego wzoru: $p \cdot v = c$, gdzie p jest ciśnieniem, v — objętością oraz c — pewną stałą. Stała c zależy od różnych czynników, na przykład od temperatury. Przypuśćmy, że zadanie polega na oszacowaniu tej stałej dla pewnego konkretnego naczynia i pewnego konkretnego gazu. Przypuśćmy dalej, że możemy zmieniać objętość naszego naczynia (na przykład za pomocą manipulowania tłokiem) i że możemy mierzyć ciśnienie, przy czym pomiar ciśnienia obarczony jest pewnym błędem przypadkowym. Jeżeli przy objętości v_1 zarejestrujemy

ciśnienie p_1 , to błąd wynosi $p_1 - \frac{c}{v_1}$. Jeżeli powtórzymy nasz pomiar n razy i przy

objętościach v_1, v_2, \dots, v_n otrzymamy wyniki p_1, p_2, \dots, p_n , to zgodnie z metodą najmniejszych kwadratów za oszacowanie stałej c należy przyjąć taką wartość, która minimalizuje wielkość

$$\left(p_1 - \frac{c}{v_1}\right)^2 + \left(p_2 - \frac{c}{v_2}\right)^2 + \dots + \left(p_n - \frac{c}{v_n}\right)^2$$

Czytelnik mógłby zauważyć, że można wykonać tylko jeden pomiar i za c przyjąć po prostu iloczyn $p_1 v_1$. Albo że można wykonać kilka razy pomiar ciśnienia przy tej samej objętości v i że wtedy należałoby minimalizować prostszą wielkość

$$\left(p_1 - \frac{c}{v}\right)^2 + \left(p_2 - \frac{c}{v}\right)^2 + \dots + \left(p_n - \frac{c}{v}\right)^2,$$

$$\frac{1}{\text{K}} = ?$$



?



Rozwiązanie zadania F 55.

W stanie równowagi energia swobodna musi mieć ekstremum. W rozważanym problemie energia swobodna jest energią powierzchniową kryształu i równa się

$$E = 2a^2\sigma_2 + 4ac\sigma_1.$$

Zgodnie z treścią zadania składnik energii swobodnej równy energii potencjalnej został pominięty.

Aby wyznaczyć stosunek a/c należy znaleźć ekstremum E przy założeniu dodatkowym, że objętość kryształu $V = a^2c$ jest stała. Innymi słowy, należy znaleźć ekstremum funkcji

$$E(a) = 2a^2\sigma_2 + 4V\sigma_1/a,$$

gdzie $V = a^2c = \text{const}$.

Obliczając pochodną i przyrównując ją do zera otrzymujemy

$$4a\sigma_2 - 4V\sigma_1/a^2 = 0$$

$$a\sigma_2 - c\sigma_1 = 0,$$

a zatem

$$\frac{a}{c} = \frac{\sigma_1}{\sigma_2}.$$

a to zadanie już rozwiązaliśmy i wiemy, że jego rozwiązaniem jest średnia arytmetyczna: $c = v(p_1 + p_2 + \dots + p_n)/n$. Czytelnik ma rację, ale okazuje się, że dokładność oszacowania c jest tym większa, im więcej wykonuje się pomiarów i im zreczniej dobiera się wartości v_1, v_2, \dots, v_n . Wyjaśnienie, dlaczego tak jest, wykracza jednak daleko poza ramy naszej krótkiej rozprawki.

Przykład 2. Na pewno każdy z Czytelników usłyszał kiedyś od lekarza, że „jak na swój wiek” jest zbyt wysoki, zbyt niski lub że ma wzrost „w normie”. Na pewno jednak nie każdy starał się dociec, co to znaczy. Spróbuję to wyjaśnić. Będzie to zarazem ilustracją pewnego dość powszechnego, ale odmiennego od poprzednich, sposobu zastosowania metody najmniejszych kwadratów. Postępuje się mianowicie w następujący sposób. „Bierze się” dużą grupę N ludzi w różnym wieku i mierzy się wzrost każdej osoby w tej grupie. Otrzymuje się wyniki: $(t_1, w_1), (t_2, w_2), \dots, (t_N, w_N)$, gdzie t_i jest wiekiem oraz w_i — wzrostem i -tej osoby. Po naniesieniu na wykres wyglądają one mniej więcej tak, jak na rys. 1 (każda kropka na tym rysunku reprezentuje jedną osobę.)

Następnie rysuje się krzywą, która ma przebiegać „mniej więcej” tak, jak układają się nasze punkty (linia ciągła na rysunku). Z kolei „wymyśla się” wzór, który byłby równaniem tej krzywej. Dla krzywych wzrostu często przyjmuje się wzór

$$(*) \quad w = \frac{a}{1 + b \cdot c^t}$$

gdzie a, b, c są pewnymi stałymi. Jest to tzw. *krzywa logistyczna*. Zgódźmy się na to równanie; od tej pory staje się ono dla nas „empirycznym prawem wzrostu”. Różnym wartościom stałych a, b, c (podobnie jak różnym wartościom stałej c w prawie Boyle’a-Mariotte’a) odpowiadają różne krzywe (rys. 2) i zadanie polega na znalezieniu takich a, b, c , które najlepiej odpowiadają wynikom naszych pomiarów. Zgodnie z naszym prawem wzrostu, wzrost osoby w wieku t_i powinien wynosić $a/(1 + b \cdot c^t)$, natomiast w wyniku pomiaru otrzymaliśmy dla tej osoby wynik w_i . Wielkość $w_i - \frac{a}{1 + b \cdot c^t}$ jest odpowiednikiem tego, co w poprzednich przykładach nazywaliśmy błędem ϵ_i . Zatem — zgodnie z metodą najmniejszych kwadratów — stałe a, b, c należy dobrać tak, żeby wielkość

$$\sum_{i=1}^N \left(w_i - \frac{a}{1 + b \cdot c^t} \right)^2$$

osiągała minimum. To zadanie jest dość skomplikowane rachunkowo i praktycznie rozwiązuje się je za pomocą komputerów (nawet dla małych wartości N). Wzór (*) z tak ustalonymi wartościami a, b, c opisuje właśnie to, co nazywa się „normalnym wzrostem w danym wieku”. Takie krzywe wzrostu można oczywiście konstruować oddzielnie dla mężczyzn i dla kobiet, czasami oddzielnie dla różnych środowisk (np. miasto — wieś), a na przykład powiedzenie, że „Szwedzi są średnio wyżsi od Polaków” oznacza, że krzywa wzrostu Szwedów leży powyżej polskiej krzywej wzrostu.

Metoda najmniejszych kwadratów jest codziennym narzędziem fizyków, inżynierów, przyrodników, ekonomistów, a jej teoria jest przedmiotem zainteresowania statystyki matematycznej.

Przykład (konkretny). Zakładamy, że wielkości X i Y łączy zależność liniowa $Y = aX + b$. W czterech pomiarach tych wielkości uzyskano wyniki

X	0	10	20	30
Y	14	15	37	41

Wyznaczyć metodą najmniejszych kwadratów przybliżenia a_0 i b_0 współczynników a i b .

Rozwiązanie. Załóżmy, ogólniej, że dysponujemy wynikami n pomiarów $(x_1, y_1), \dots, (x_n, y_n)$. Należy wyznaczyć liczby a_0 i b_0 tak, by wyrażenie

$$W(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2$$

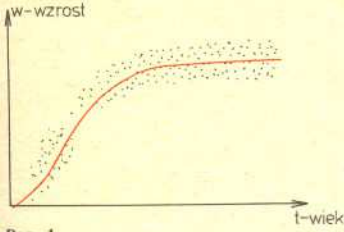
osiągnęło wartość najmniejszą dla $a = a_0$ i $b = b_0$. Ponieważ funkcja ta ma tylko jedno minimum, to można je wyznaczyć elementarnie w sposób następujący. Funkcja $W(a_0, b)$ jest trójmianem kwadratowym względem b :

$$W(a_0, b) = nb^2 - \left[2 \sum_i (y_i - a_0 x_i) \right] b + \sum_i (y_i - a_0 x_i)^2$$

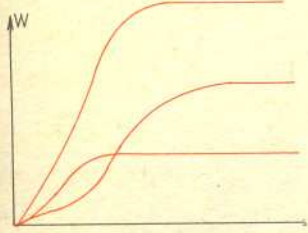
i minimum osiąga w punkcie $b = \frac{1}{2} \frac{2 \sum_i (y_i - a_0 x_i)}{n}$, który musi być identyczny z b_0 , skąd

$$(*) \quad b_0 = \bar{y} - a_0 \bar{x}$$

gdzie \bar{x}, \bar{y} są odpowiednio średnimi arytmetycznymi liczb x_1, \dots, x_n i y_1, \dots, y_n .



Rys. 1



Rys. 2



Podobnie, minimum trójmianu kwadratowego względem a

$$W(a, b_0) = \left[\sum_i x_i^2 \right] a^2 - \left[2 \sum_i (y_i - b_0) x_i \right] a + \sum_i (y_i - b_0)^2$$

osiągane jest w punkcie
$$a_0 = \frac{1}{2} \frac{2 \sum_i (y_i - b_0) x_i}{\sum_i x_i^2},$$
 skąd

$$(**) \quad a_0 \sum_i x_i^2 = \sum_i x_i y_i - b_0 \sum_i x_i.$$

Podstawiając (*) do (**) otrzymamy

$$a_0 \left[\sum_i x_i^2 - \bar{x} \sum_i x_i \right] = \sum_i x_i y_i - \bar{y} \sum_i x_i,$$

skąd po przekształceniach

$$(***) \quad a_0 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}.$$

Wzory (***) i (*) dają szukane rozwiązanie. W naszym konkretnym przykładzie otrzymamy

$$a_0 = 1,03, \quad b_0 = 11,3.$$

Dla znających rachunek prawdopodobieństwa: wystarczy skorzystać ze wzorów $E[(X-EX)^2] = E(X^2) - [E(X)]^2$ i $E[(X-EX)(Y-EY)] = E(XY) - EY \cdot EX$.

Zasady ekstremalne w termodynamice

Dr Bogdan CICHOCKI

Zanim przejdziemy do omówienia zagadnienia wspomnianego w tytule niniejszego artykułu, przypomnimy podstawowe pojęcia termodynamiki. Wiemy na podstawie doświadczenia, że każde ciało makroskopowe, niezależnie od stanu początkowego, osiąga po pewnym czasie stan równowagi. Stan taki można scharakteryzować przez podanie niewielkiej liczby parametrów. Przykładowo, stan równowagi określonej liczby moli gazu jednoskładnikowego opisują w pełni dwie wielkości (np. ciśnienie p i objętość V). Na skutek działania czynników zewnętrznych układ fizyczny może przechodzić z jednego stanu równowagi do innego. Przejściem takim zwanym *procesem termodynamicznym* rządzą pewne prawa. Zgodnie z treścią pierwszej zasady termodynamiki w procesie termodynamicznym zmiana energii wewnętrznej U układu (tzn. różnica między wartościami U w stanach końcowym i początkowym) jest równa sumie ciepła Q dostarczonego układowi i pracy L wykonanej nad układem, czyli

$$(1) \quad U_k - U_p = Q + L.$$

Jeżeli ciało fizyczne składa się z dwóch części, powiedzmy 1 i 2, to jego energia wewnętrzna U jest równa

$$(2) \quad U = U_1 + U_2 + U_{12},$$

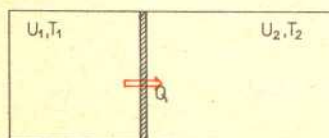
gdzie U_1, U_2 są energiami wewnętrznymi poszczególnych części, zaś U_{12} energią oddziaływania tych części. Ponieważ dla ciał makroskopowych $U_1 \gg U_{12}$ i $U_2 \gg U_{12}$, więc w przybliżeniu zachodzi związek

$$(3) \quad U = U_1 + U_2.$$

Mówimy, że wielkość U ma własność addytywności.

Pierwsza zasada termodynamiki nie charakteryzuje w pełni procesów występujących w przyrodzie. Aby się o tym przekonać, rozważmy następujący przykład. Wyobraźmy sobie układ fizyczny składający się z dwóch podukładów 1 i 2, o energiach wewnętrznych i temperaturach odpowiednio U_1, T_1 i U_2, T_2 . Niech między tymi podukładami znajduje się ścianka przewodząca ciepło. Z doświadczenia wiemy, że układ taki znajduje się w stanie równowagi, o ile $T_1 = T_2$. Jeżeli tak nie jest, to z układu o wyższej temperaturze przepływa ciepło do układu o niższej temperaturze, aż do momentu wyrównania się temperatur. Zgodnie z pierwszą zasadą każdy proces, dla którego zachodzi równość (1), jest możliwy. Również taki, w którym ciepło przepływa w kierunku przeciwnym niż ten, o którym wspomnieliśmy powyżej. Wynika z tego, że pierwsza zasada wymaga uzupełnienia. Zadanie to spełnia druga zasada termodynamiki.

Istnieje wiele równoważnych sformułowań drugiej zasady termodynamiki. Właściwie to już podaliśmy jedno z nich. Mianowicie, jest to twierdzenie dotyczące kierunku przepływu ciepła. Sformułowanie to podane po raz pierwszy przez Clausiusa jest bardzo pogładowe, jednakże droga prowadząca od niego do wniosków o charakterze ogólnym jest dość zawiła. Omówimy teraz inne sformułowanie drugiej zasady termodynamiki, które przybiera postać zasady



$T_1 > T_2$