

Analiza niepewności pomiarowych

Wstęp do analizy danych

Wykład 6

Metoda najmniejszych kwadratów i jej zastosowania:
średnia ważona, proporcjonalność, prosta.

1 Uwagi wstępne

Dotychczas zajmowaliśmy się wyznaczaniem wartości pojedynczych wielkości fizycznych. Analizowaliśmy przypadki, gdy możliwy był bezpośredni pomiar interesującej nas wielkości lub gdy mogliśmy jej wartość obliczyć korzystając ze znanej nam funkcji, której argumentami są wielkości dostępne bezpośrednim pomiarom. Postać funkcji mogła wynikać z rozważań teoretycznych lub wcześniejszych badań doświadczalnych. Zajmijmy się teraz analizą procesu doświadczalnego “odkrywania” związku między dwoma wielkościami fizycznymi. Jako przykład weźmy badanie zmian długości l stalowego pręta pod wpływem zmian temperatury T . Doświadczenie mogłoby wyglądać tak: umieszczamy pręt w kąpeli olejowej, której temperaturę potrafimy dokładnie kontrolować i powoli zmieniać i po każdej zmianie temperatury mierzymy długość pręta. Otrzymujemy zbiór par liczb: $(T_i, l_i), i = 0, \dots, N$, przy czym dane dla $i = 0$ odnoszą się do początkowych wartości temperatury i długości. Ponieważ kontrolujemy wartość temperatury, to pomiary możemy wykonać wielokrotnie mierząc długość dla tych samych temperatur T_i podczas kilkakrotnego ogrzewania i chłodzenia pręta. Dla każdej z temperatur T_i obliczamy średnią wartość długości \bar{l}_i oraz wyznaczamy niepewność $u_i := u(\bar{l}_i)$ tej średniej. Wydaje się naturalne, że nasze wyniki będziemy analizowali przyjmując, że długość l jest funkcją temperatury T . Postać tej funkcji możemy próbować odgadnąć wykonując wykres $\bar{l}_i := l(T_i)$. Spodziewamy się, że punkty wykresu ułożą się wzdłuż prostej:

$$l = l_0(1 + \beta(T - T_0)),$$

bo taką zależność obserwowano wielokrotnie w przypadku badania prętów z różnych metali i ich stopów, gdy różnica $T - T_0$ jest niezbyt duża, a obie temperatury są niższe od temperatury topnienia materiału pręta. Pozostaje nam wyznaczenie wartości parametrów l_0 i β . Biorąc temperaturę za zmienną niezależną założyliśmy, że tę właśnie wielkość nie tylko potrafimy kontrolować, ale także zmierzyć dokładnie. Poszukujemy odpowiedzi na pytanie, jak na podstawie uzyskanych wyników pomiarów:

$$T_i \longmapsto \bar{l}_i \pm u(\bar{l}_i), \quad i = 0, \dots, N$$

wyznaczyć *najlepsze* wartości parametrów l_0 i β funkcji $l = l_0(1 + \beta(T - T_0))$ opisującej związek dwóch wielkości fizycznych T i l . Nasza odpowiedź poza oceną wartości parametrów powinna

pozwolić ocenić także ich niepewności.

W §19 *slajdów do wykładu* przedstawione jest rozwiązanie ogólniejszego zagadnienia, w którym funkcja opisująca związek dwóch wielkości zależy od kilku parametrów. Ogólna metoda jest następnie zastosowana do kilku najprostszych, często spotykanych w praktyce postaci funkcji: w §20 funkcji stałej, w §21 proporcjonalności oraz “prostej” w §22, czyli zależności takiej, jak w przytoczonym powyżej przykładzie.

Warto w tym miejscu podkreślić wagę wykonania wykresu otrzymanych wyników. Jeśli, jak w naszym przykładzie, znajomość wyników pomiarów podobnych do naszych daje nam wyraźną wskazówkę co do postaci badanej zależności $l(T)$, to wykres naszych danych potwierdzi poprawność wyboru postaci funkcji. Po wyznaczeniu wartości parametrów zawsze warto na wykres danych doświadczalnych nanieść także wykres wybranej funkcji z podstawionymi wyznaczonymi wartościami parametrów. To najprostszy test poprawności naszego dopasowania funkcji do danych doświadczalnych. Jest on szczególnie skuteczny jako metoda “wyłapywania” błędów grubych w naszym postępowaniu, np. błędów rachunkowych (każdemu mogą się przydarzyć takie pomyłki). W dalszej części wykładu zajmiemy się ilościowymi testami zgodności dopasowanej funkcji z danymi doświadczalnymi.

Dodatkowo, jeśli wykonujemy pomiary dla dużego zakresu wartości zmiennej niezależnej x , to wykres wyników pomiarów może pozwolić określić zakres stosowalności wzoru, który “dopasujemy” do danych. Dla przykładu: jeśli badamy zmiany długości l pręta pod wpływem rozciągającej go siły F (spodziewamy się zachowania zgodnego z prawem Hooke’a), to odstępstwo od liniowości związku $l(F)$ oznacza wejście w obszar nieodwracalnych odkształceń plastycznych.

Uwaga: W *slajdach do wykładu* zamiast u_i posługuję się oznaczeniem s_i , które zwykle oznacza *statystyczny* wkład do niepewności – tj. wyznaczony na podstawie rozrzutu wyników serii pomiarów – inaczej mówiąc s_i oznacza niepewność pomiaru wyznaczoną **metodą A** według terminologii przyjętej przez BIPM (patrz przewodnik ze strony Głównego Urzędu Miar cytowany w **Wykładzie 4**). Tutaj stosuję ogólniejszą notację u_i odnoszącą się do pełnej oceny niepewności **metodą B**.

2 Metoda najmniejszych kwadratów

W §19 *slajdów do wykładu* formułujemy metodę znajdowania odpowiedzi na postawione powyżej pytanie. Niech x oznacza zmienną niezależną (argument funkcji), a y zmienną zależną.

Zakładamy, że wybraliśmy postać funkcji opisującej wyniki naszych pomiarów:

$y = f(x; a_1, a_2, \dots, a_k)$. Funkcja ta, poza zmienną x , zawiera także k parametrów a_j – np. f może być wielomianem stopnia $k - 1$ zmiennej x , a parametry a_0, \dots, a_{k-1} współczynnikami tego wielomianu. Zakładamy dalej, że dla $i = 1, \dots, N > k$ różnych, *dokładnie* znanych wartości x_i wykonaliśmy pomiary odpowiadających im wielkości y i w wyniku tych pomiarów wyznaczyliśmy $\hat{y}_i \pm u(\hat{y}_i)$ – dla uproszczenia zapisu dalej przyjmijmy oznaczenie $u_i := u(\hat{y}_i)$:

$$x_i \longmapsto \hat{y}_i \pm u_i; i = 1, \dots, N.$$

Jak zwykle zakładamy, że wyeliminowane zostały błędy grube i systematyczne, a więc:

$$\mathcal{E}(\hat{y}_i) = f(x_i; a_1, \dots, a_k),$$

$$\mathcal{E}(u_i^2) = \sigma_i^2,$$

gdzie σ_i^2 oznacza wariancję rozkładu pomiarów wielkości y przy ustalonej wartości $x = x_i$, a $f(x_i; a_1, \dots, a_k)$ *dokładną* wartość y dla ustalonego x_i . Zakładamy także, że \hat{y}_i są niezależnymi zmiennymi losowymi i każde \hat{y}_i podlega rozkładowi Gaussa o średniej $f(x_i; a_1, \dots, a_k)$ i dyspersji σ_i . Wartości parametrów a_i nie znamy, ale naturalne jest przyjęcie, że wyniki przeprowadzonych przez nas pomiarów bliskie są wartościom najbardziej prawdopodobnym (typowym). Pozwoli nam to na sformułowanie metody poszukiwania *najlepszych* wartości parametrów a_i . Zgodnie z naszymi założeniami N niezależnych zmiennych losowych \hat{y}_i (\hat{y}_i – wynik pomiaru przy ustalonym x_i) podlega N -wymiarowemu rozkładowi Gaussa:

$$g(\hat{y}_1, \dots, \hat{y}_N; a_1, \dots, a_k) = \frac{1}{(2\pi)^{N/2} \sigma_1 \cdots \sigma_N} \exp\left(-\sum_{i=1}^N \frac{(\hat{y}_i - f(x_i; a_1, \dots, a_k))^2}{2\sigma_i^2}\right),$$

zależnemu od k parametrów a_i , wielkości σ_i^2 są liczbami równymi wariancjom wielkości \hat{y}_i .

Stwierdzenie: nasze wyniki pomiarów, \hat{y}_i odpowiadają sytuacji typowej – są bliskie maksimum gęstości prawdopodobieństwa g opisanej powyższym wzorem – oznacza, że rzeczywiste wartości parametrów a_i są takie, że g przyjmuje wartość maksymalną, gdy podstawimy wartości $\hat{y}_1, \dots, \hat{y}_N$.

Wniosek: najlepsze oceny wartości parametrów a_1, \dots, a_k odpowiadają maksimum g przy ustalonych $\hat{y}_1, \dots, \hat{y}_N$. Każda z ocen wartości a_1, \dots, a_k będzie funkcją wartości *wszystkich* $\hat{y}_1, \dots, \hat{y}_N$. Niepewności ocen a_i wyznaczymy posługując się prawem propagacji niepewności na podstawie niepewności u_i , które we wzorach końcowych podstawimy w miejsce σ_i .

Poszukiwanie maksimum g jest równoważne poszukiwaniu maksimum:

$$\ln(g) = -\sum_{i=1}^N \frac{(\hat{y}_i - f(x_i; a_1, \dots, a_k))^2}{2\sigma_i^2} - \frac{N}{2} \ln(2\pi) - \sum_{i=1}^N \ln(\sigma_i).$$

Wartości \hat{y}_i oraz σ_i nie zależą od a_1, \dots, a_k i ostatecznie poszukiwanie maksimum gęstości g sprowadza się do poszukiwania **minimum** funkcji:

$$\mathcal{R}(a_1, \dots, a_k) := \sum_{i=1}^N \frac{(\hat{y}_i - f(x_i; a_1, \dots, a_k))^2}{\sigma_i^2} \longrightarrow \text{minimum}.$$

względem a_1, \dots, a_k . Po wyprowadzeniu wzorów, w obliczeniach, w miejsce $\sigma_1, \dots, \sigma_N$ podstawiamy u_i , tj. nasze oceny tych wartości.

3 Średnia ważona

W poprzednich wykładach szczegółowo omawiany był bezpośredni pomiar pojedynczej wielkości fizycznej. Analizowaliśmy tam wyniki pomiarów powtarzanych wielokrotnie zawsze tą samą metodą, w tych samych warunkach. Wyznaczaliśmy na ich podstawie ocenę wartości mierzonej wielkości \hat{x} i jej niepewność $u(\hat{x})$. Wyobraźmy sobie, że podobne pomiary tej samej wielkości wykonywali też nasi koledzy i koleżanki. Oczywiście, wyniki otrzymane przez poszczególne osoby różnią się między sobą zarówno, jeśli chodzi o wartość mierzonej wielkości, jak i o jej niepewność. Pomiary mogły też być wykonywane różnymi metodami.

Jak na podstawie tych wyników wyznaczyć najlepszą ocenę mierzonej wielkości?

Dysponujemy N wynikami pomiarów tej samej wielkości x :

$$x = \hat{x}_i \pm u_i ; i = 1, \dots, N.$$

W *slajdach do wykładu* zamiast u_i posługuję się oznaczeniem s_i , które zwykle oznacza *statystyczny wkład do niepewności* – tj. wyznaczony na podstawie rozrzutu wyników serii pomiarów.

Posłużymy się metodą najmniejszych kwadratów:

- Zmienną niezależną jest numer pomiaru i .
- “Dopasowywaną” funkcją i jest **funkcja stała**: $f(i) := \mu$; niech \hat{x} oznacza poszukiwaną, końcową ocenę wartości μ na podstawie wszystkich wyników $\hat{x}_i \pm u_i, i = 1, \dots, N$.
- Zakładamy, że dla wszystkich i , $\mathcal{E}(\hat{x}_i) = \mu$ – dokładnej wartości x oraz $\mathcal{E}(u_i^2) = \sigma_i^2$ – dokładnej wartości wariancji zmiennej losowej \hat{x}_i .
- Poszukujemy minimum

$$\mathcal{R}(\hat{x}) = \sum_{i=1}^N \frac{(\hat{x}_i - \hat{x})^2}{\sigma_i^2},$$

gdzie liczby σ_i^2 oznaczają wariancje zmiennych losowych \hat{x}_i .

- W obliczeniach, we wzorach końcowych, w miejsce nieznanymi wartości σ_i podstawimy znane nam ich oceny u_i .

Po zróżniczkowaniu $\mathcal{R}(\hat{x})$ względem \hat{x} i przyrównaniu pochodnej do zera otrzymujemy równanie na \hat{x}_w , wartość \hat{x} minimalizującą $\mathcal{R}(\hat{x})$:

$$-2 \sum_{i=1}^N \frac{(\hat{x}_i - \hat{x}_w)}{\sigma_i^2} = 0.$$

Rozwiązaniem tego równania jest:

$$\hat{x}_w = \frac{\sum_{i=1}^N \hat{x}_i / \sigma_i^2}{\sum_{i=1}^N 1 / \sigma_i^2}.$$

Otrzymana zmienna losowa \hat{x}_w jest liniową funkcją wszystkich zmiennych losowych (wyników pomiarów) \hat{x}_i , a każda ze zmiennych \hat{x}_i mnożona jest przez współczynnik:

$$\hat{x}_w = \sum_{i=1}^N d_i \hat{x}_i; \quad d_i = \frac{1 / \sigma_i^2}{\sum_{j=1}^N 1 / \sigma_j^2}$$

Przy czym, jak łatwo sprawdzić:

$$\sum_{i=1}^N d_i = 1.$$

Podstawienie współczynników d_i do prawa *propagacji niepewności* (str. 39 *slajdów do wykładu*), po kilku przekształceniach, prowadzi do podanej oceny wariancji zmiennej losowej \hat{x}_w :

$$u_{int}^2(\hat{x}_w) := \mathcal{E}((\hat{x}_w - \mu)^2) = \frac{1}{\sum_{i=1}^N \frac{1}{\sigma_i^2}}.$$

Warto zwrócić uwagę na fakt, że wyrażenie na u_{int}^2 zawiera wyłącznie wartości σ_i^2 , a nie zawiera samych \hat{x}_i . Tym samym, u_{int}^2 nie jest zmienną losową i właściwsze byłoby oznaczenie σ_{int}^2 , ze względu jednak na fakt, że otrzymane wyrażenie będzie dalej wykorzystywane do wyznaczania *niepewności* \hat{x}_w używam oznaczenia u_{int}^2 .

Odpowiedź 1 (str. 44 *slajdów do wykładu*): $x = \hat{x}_w \pm u_{int}(\hat{x}_w)$

podsumowuje uzyskane wyżej wyniki. Indeks “w” w \hat{x}_w zwraca uwagę na fakt, że otrzymane wyrażenie jest **średnią ważoną** wyników \hat{x}_i z wagami $1/\sigma_i^2$.

Łatwo sprawdzić, że gdy **wszystkie wagi są równe (tzn. $\sigma_i = \sigma$)**, to:

$$\begin{aligned}\hat{x}_w &= \frac{1}{N} \sum_{i=1}^N \hat{x}_i, \\ u_{int}^2 &= \frac{\sigma^2}{N},\end{aligned}$$

zgodnie z wcześniejszym wnioskiem, że wariancja średniej arytmetycznej próbki N elementów losowanych z tego samego rozkładu równa jest $1/N$ wariancji rozkładu (str. 26 *slajdów do wykładu*).

Dla lepszego zrozumienia otrzymanych wyżej wyników sprawdźmy ich zastosowanie do przypadku, gdy \hat{x}_i oraz $u_i, i = 1, \dots, N$ otrzymaliśmy w następujący sposób:

- wielokrotnie powtarzano pomiary tej samej wielkości fizycznej stosując tę samą metodę i te same przyrządy pomiarowe – uzyskano bardzo długą serię niezależnych wyników (np. zmierzono 216 razy czas trwania pojedynczego okresu wahadła);
- otrzymane wyniki podzielono na N krótszych podserii (np. serię 216 pomiarów okresu podzielono na $N = 7$ podserii o liczebnościach: $n_1 = 108, n_2 = 54, n_3 = 18, n_4 = 9, n_5 = 9, n_6 = 9, n_7 = 9$);
- dla każdej z podserii wyznaczono jej średnią: $\hat{x}_i, i = 1, \dots, N$;
- rozrzut wyników w każdej z podserii wskazuje, że wkład błędu przypadkowego do niepewności każdej z wartości \hat{x}_i wyraźnie dominuje nad wkładem ograniczonej dokładności przyrządów, a więc $u_i \approx s_i \approx \sigma_i, i = 1, \dots, N$;
- pomiary były wykonywane tą samą metodą i za pomocą tych samych przyrządów, a więc wariancja σ_i^2 zmiennej \hat{x}_i – średniej n_i pomiarów, wynosi $\sigma_i^2 = \sigma^2/n_i$.

Czy znając jedynie $\hat{x}_i, \sigma_i, i = 1, \dots, N$ potrafimy uzyskać poprawną ocenę wartości \hat{x} – średniej wszystkich wyników długiej serii oraz poprawną ocenę niepewności \hat{x} ?

Model alternatywny (str. 44 *slajdów do wykładu*) daje odpowiedź na to pytanie i przedstawia rozumowanie prowadzące do wyznaczenia wartości \hat{x} i jej wariancji $u_{\hat{x}}^2$.

Odpowiedź 2. (str. 46 *slajdów do wykładu*): $x = \hat{x}_w \pm u_{ext}$

$$u_{ext}^2 := \frac{\sum_{i=1}^N [(\hat{x}_i - \hat{x}_w)^2 / \sigma_i^2]}{(N-1) \sum_{i=1}^N (1/\sigma_i^2)} = \frac{\sum_{i=1}^N [(\hat{x}_i - \hat{x}_w)^2 / \sigma_i^2]}{(N-1)} u_{int}^2.$$

Otrzymana ocena \hat{x} (przypominam, że \hat{x} jest oceną wartości μ na podstawie znanych wyników \hat{x}_i) jest średnią ważoną wszystkich \hat{x}_i , dokładnie taką, jak uzyskana poprzez zastosowanie metody najmniejszych kwadratów: $\hat{x} = \hat{x}_w$. Ocena wariancji \hat{x} , równa u_{ext}^2 , jest jednak w tym przypadku inna. W przeciwieństwie do u_{int}^2 wzór na u_{ext}^2 zawiera nie tylko wartości σ_i , ale także sumę kwadratów odchyłeń \hat{x}_i od \hat{x}_w .

Łatwo sprawdzić, że gdy **wszystkie σ_i są jednakowe (tzn. $\sigma_i = \sigma$)**, to:

$$\begin{aligned} \hat{x}_w &= \frac{1}{N} \sum_{i=1}^N \hat{x}_i, \\ u_{ext}^2 &= \frac{1}{N(N-1)} \sum_{i=1}^N (\hat{x}_i - \hat{x}_w)^2 \end{aligned}$$

zgodnie z wcześniej wyprowadzonym wzorem na ocenę wariancji rozkładu na podstawie próbki o liczebności N (str. 44 *slajdów do wykładu*).

Wróćmy do przypadku, gdy σ_i są różne. Wiemy, że dla $\hat{x} = \hat{x}_w$ wyrażenie $\mathcal{R}(\hat{x})$ przyjmuje wartość minimalną. Wyrażenie to jest funkcją zmiennych losowych \hat{x}_i , a więc też jest zmienną losową. Znajdźmy jego wartość oczekiwaną:

$$\begin{aligned} \mathcal{E}(\mathcal{R}(\hat{x}_w)) &= \mathcal{E}\left(\sum_{i=1}^N \frac{(\hat{x}_i - \hat{x}_w)^2}{\sigma_i^2}\right) \\ &= \sum_{i=1}^N \frac{1}{\sigma_i^2} \mathcal{E}((\hat{x}_i - \mu + \mu - \hat{x}_w)^2) \\ &= \sum_{i=1}^N \frac{1}{\sigma_i^2} \mathcal{E}((\hat{x}_i - \mu)^2 + (\hat{x}_w - \mu)^2 - 2(\hat{x}_i - \mu)(\hat{x}_w - \mu)) \\ &= N - 1. \end{aligned}$$

Ostatnią równość otrzymaliśmy po skorzystaniu z założeń, że zmienne \hat{x}_i i \hat{x}_j dla $i \neq j$ są niezależne, $\hat{x}_w = \sum_{i=1}^N d_i \hat{x}_i$ (z d_i zdefiniowanym wyżej), $\mathcal{E}(\hat{x}_i) = \mu$ oraz $\mathcal{E}(\hat{x}_i - \mu)^2 = \sigma_i^2$.

Wniosek: Wartość oczekiwana u_{ext}^2 jest równa u_{int}^2 .

Po wyznaczeniu \hat{x}_w obliczmy wartości u_{int} i u_{ext} podstawiając wartości u_i w miejsce σ_i do wzorów definiujących u_{int} i u_{ext} . Spodziewamy się, że otrzymane wartości będą niewiele się różniły.

Jako ocenę niepewności \hat{x}_w bierzemy większą z wartości u_{int} i u_{ext} .

Uwaga: Ścisłe biorąc, jeśli u_{ext} jest znacznie większe od u_{int} , oznacza to, że uśredniane wartości \hat{x}_i nie są ze sobą zgodne – niektóre z nich za bardzo różnią się od zakładanej wspólnej wartości oczekiwanej μ . Może to być wynik wystąpienia błędów systematycznych w pomiarach niektórych z nich. Podanie ostatecznej oceny wymagałoby więc dokładnego przeanalizowania sposobu pomiaru każdej z wartości \hat{x}_i w celu znalezienia źródła niezgodności. Często jednak taka analiza nie jest możliwa, gdy np. uśredniane wartości otrzymaliśmy z różnych laboratoriów. Wtedy obliczenie średniej ważonej i przyjęcie większej z obliczonych niepewności za niepewność tej średniej jest rozwiązaniem najrozsądniejszym i często praktykowanym w braku lepszego. Jeśli jednak krytyczna analiza źródłowych danych jest możliwa (np. mamy dostęp do pełnej dokumentacji przebiegu pomiarów każdej z wielkości \hat{x}_i) to należy taką analizę przeprowadzić w celu znalezienia przyczyny rozbieżności danych. Jednym z pięknych przykładów zasadności tego typu analizy jest historia odkrycia argonu przez Johna Williama Strutta (Lorda Rayleigh) i Williama Ramsay’a. Wykonywane przez tych uczonych pomiary masy cząsteczkowej azotu prowadziły do skupiania wyników pomiarów wokół dwóch nieznacznie różnych wartości. Dokładna analiza wykazała, że azot pochodzący z reakcji chemicznych rozkładu jego związków prowadził do oceny masy o 0,5% mniejszej niż w przypadku azotu z powietrza. W dalszych badaniach Rayleigh i Ramsay wykazali, że źródłem tej różnicy jest obecność w powietrzu gazu “nieczynnego” chemicznie i tak odkryli gaz szlachetny – argon. W roku 1904 otrzymali za to odkrycie Nagrody Nobla: Rayleigh z fizyki, a Ramsay z chemii. Jak widać staranność w analizowaniu danych jest warta zachodu.

Rysunek 7 *slajdów do wykładu* przedstawia przykładowe modele rozkładów pradowpodobieństwa odpowiadające pomiarom, dla których $u_{int} > u_{ext}$, a Rysunek 8 analogiczne rozkłady w przypadku $u_{int} < u_{ext}$ (pomyśl takiego porównania zaczerpnąłem z książki: W. H. Gränicher, *Messung beendet – was nun?*, B. G. Teubner, Stuttgart, 1994).

4 Wyznaczanie współczynnika proporcjonalności

§21 *slajdów do wykładu* poświęcony jest znajdowaniu współczynnika proporcjonalności a między zmienną niezależną x i zmienną zależną y : $y = ax$. Zakładamy, że potrafimy ustalać *dokładne* wartości zmiennej x , dla których mierzymy odpowiadające im wartości y . Pomiary wykonujemy dla N różnych wartości x :

$$x_i \mapsto \hat{y}_i \pm u_i, i = 1, \dots, N.$$

Zakładamy także, że pomiary y_i i y_j są niezależnie dla $i \neq j$, a między wielkościami x i y zachodzi związek $y = ax$. Postępowanie prowadzące do znalezienia najlepszej oceny wartości parametru a przebiega całkowicie analogicznie do rozumowania prowadzącego do średniej ważonej:

- Traktujemy każdy z wyników pomiarów \hat{y}_i jak zmienną losową o wariancji σ_i^2 .
- Poszukujemy minimum funkcji:

$$\mathcal{R}(a) = \sum_{i=1}^N \frac{(\hat{y}_i - ax_i)^2}{\sigma_i^2}$$

względem a .

Przyrównanie do zera pochodnej $\mathcal{R}(a)$ względem a prowadzi do warunku:

$$\hat{a} \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} = \sum_{i=1}^N \frac{\hat{y}_i x_i}{\sigma_i^2},$$

gdzie \hat{a} jest poszukiwaną oceną wartości parametru a . Otrzymujemy:

$$\hat{a} = \frac{\sum_{i=1}^N \hat{y}_i x_i / \sigma_i^2}{\sum_{i=1}^N x_i^2 / \sigma_i^2}.$$

Ocena \hat{a} jest liniową funkcją zmiennych losowych \hat{y}_i (wyników pomiarów). Możemy zastosować wzór na propagację niepewności i otrzymać wzór na wariancję zmiennej losowej \hat{a} :

$$u^2(\hat{a}) = \frac{1}{\sum_{i=1}^N x_i^2 / \sigma_i^2}.$$

Podobnie, jak w przypadku u_{int}^2 obliczone wyżej $u^2(\hat{a})$ nie jest zmienną losową – używam oznaczenia $u^2(\hat{a})$, bo otrzymany wzór będzie służył do wyznaczania *niepewności* \hat{a} .

Jak zwykle, wykonując obliczenia w miejsce σ_i^2 podstawiamy oceny wartości $u_i^2 := u^2(\hat{y}_i)$.

W dalszych rozważaniach bardzo przydatne będzie twierdzenie dotyczące wartości oczekiwanej $\mathcal{R}(\hat{a}_1, \dots, \hat{a}_k)$ w przypadku, gdy metodą najmniejszych kwadratów znaleźliśmy oceny wartości parametrów $\hat{a}_1, \dots, \hat{a}_k$ funkcji $y = f(x; a_1, \dots, a_k)$ opisującej zależność zmiennej losowej y od argumentu x funkcji f :

Twierdzenie: Jeżeli $f(x; a_1, \dots, a_k)$ jest **liniową funkcją** parametrów a_1, \dots, a_k , wartości parametrów $\hat{a}_1, \dots, \hat{a}_k$ zostały znalezione za pomocą dopasowania metodą najmniejszych kwadratów zależności $y_i = f(x_i, a_1, \dots, a_k)$ do zbioru wartości zmiennych losowych \hat{y}_i o wariancjach σ_i^2 to:

$$\mathcal{E}(\mathcal{R}(\hat{a}_1, \dots, \hat{a}_k)) := \mathcal{E} \left(\sum_{i=1}^N \frac{(\hat{y}_i - f(x_i; \hat{a}_1, \dots, \hat{a}_k))^2}{\sigma_i^2} \right) = N - k,$$

gdzie N jest liczbą zmiennych losowych \hat{y}_i , a k liczbą *dopasowanych* parametrów.

Jeśli znamy jedynie wyniki pomiarów \hat{y}_i , a nie znamy ich niepewności, ale mamy podstawy zakładać, że wszystkie te niepewności mają tę samą wartość (lub wartości bardzo zbliżone), to powyższe twierdzenie pozwala oszacować tę niepewność.

W przypadku wyznaczania współczynnika proporcjonalności, równość wszystkich σ_i prowadzi do uproszczenia wzoru na \hat{a} do postaci:

$$\hat{a} = \frac{\sum_{i=1}^N \hat{y}_i x_i}{\sum_{i=1}^N x_i^2}.$$

Skorzystanie z przytoczonego powyżej twierdzenia pozwala oszacować wspólną dla wszystkich \hat{y}_i wartość σ . Odpowiednią oceną wartości σ^2 jest u^2 zdefiniowane jako:

$$u^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \hat{a}x_i)^2}{N-1}$$

Mając \hat{a} i ocenę σ^2 możemy oszacować niepewność \hat{a} (podstawiamy u^2 w miejsce σ^2):

$$u^2(\hat{a}) = \frac{u^2}{\sum_{i=1}^N x_i^2} = \frac{\sum_{i=1}^N (\hat{y}_i - \hat{a}x_i)^2}{(N-1) \sum_{i=1}^N x_i^2}.$$

Uwaga (dla dociekliwych): W naszych rozważaniach \hat{y}_i są zmiennymi losowymi, a x_i *dokładnie znanymi* wartościami, można by więc zaproponować inną metodę wyznaczania współczynnika proporcjonalności a . Jeśli zależność $y = ax$ jest spełniona, to dla każdej pary x_i, \hat{y}_i można obliczyć $\hat{a}_i = \hat{y}_i/x_i$, a wyniki uśrednić. To dobry pomysł, trzeba jednak pamiętać, że niepewność tak obliczonego \hat{a}_i wynosi $u(\hat{a}_i) = u(\hat{y}_i)/x_i = u_i/x_i \approx \sigma_i/x_i$ i nawet przy jednakowych $u(\hat{y}_i)$ obliczać średnią ważoną.

Jak łatwo sprawdzić wynik będzie identyczny z wyprowadzonym powyżej:

$$\hat{a} = \frac{\sum_{i=1}^N \hat{y}_i x_i / \sigma_i^2}{\sum_{i=1}^N x_i^2 / \sigma_i^2}.$$

Obliczmy teraz u_{int}^2 średniej ważonej \hat{a}_i . Otrzymujemy:

$$u_{int}^2(\hat{a}) = \frac{1}{\sum_{i=1}^N x_i^2 / \sigma_i^2}.$$

Jest to dokładnie ten sam wzór jaki przed chwilą wyprowadziliśmy dla niepewności \hat{a} .

Nic nie stoi jednak na przeszkodzie, żeby obliczyć także odpowiednik u_{ext}^2 :

$$u_{ext}^2(\hat{a}) = \frac{\sum_{i=1}^N (\hat{a}_i - \hat{a})^2 x_i^2 / \sigma_i^2}{(N-1) \sum_{j=1}^N x_j^2 / \sigma_j^2} = \frac{\sum_{i=1}^N (\hat{y}_i - \hat{a}x_i)^2 / \sigma_i^2}{(N-1) \sum_{j=1}^N x_j^2 / \sigma_j^2} = \frac{\sum_{i=1}^N (\hat{y}_i - \hat{a}x_i)^2 / \sigma_i^2}{N-1} u_{int}^2(\hat{a}).$$

Nasz wynik: $u_{ext}^2(\hat{a}) = Au_{int}^2(\hat{a})$ ze współczynnikiem A , o którym, na mocy przytoczonego powyżej twierdzenia, wiemy, że $\mathcal{E}(A) = 1$. Po podstawieniu niepewności u_i wyników pomiarów \hat{y}_i w miejsce σ_i we wzorach na \hat{a} oraz $u_{ext}^2(\hat{a})$ i $u_{int}^2(\hat{a})$, otrzymamy wartość współczynnika A dla naszych danych. Jeśli ta wartość będzie znacznie większa od 1 (dokładnym określeniem, co w tym przypadku oznacza "znacznie większa" zajmiemy się w następnym wykładzie), to będzie znaczyło, że nasze pomiary są niezgodne z założeniem: $y = ax$ i powinniśmy to założenie porzucić. Tak postąpiłby matematyk, po uprzednim doprecyzowaniu kryterium ilościowego. Życie fizyka nie jest jednak aż tak proste, jak to sobie wyobrażają matematycy. Zdarzają się sytuacje, w których poważne przesłanki teoretyczne i wyniki doświadczeń podobnych do naszych nie pozwalają nam odrzucić hipotezy $y = ax$, a znajomość wartości parametru a jest nam

niezbędna. Rozsądnym wyjściem jest wtedy zaakceptowanie obliczonej wartości \hat{a} i przyjęcie za niepewność tego wyniku obliczonej wartości $u_{ext}(\hat{a})$. Taką decyzję można interpretować jako uznanie, że niedoceniłiśmy niepewności naszych pomiarów. Przyjęcie $u_{ext}(\hat{a})$ za niepewność \hat{a} jest równoważne przemnożeniu wszystkich niepewności u_i przez ten sam czynnik $u_i \mapsto u_i \sqrt{A}$. Taka filozofia jest w fizyce często stosowana od czasu publikacji: Raymond T. Birge, *Physical Review* **40**, 207 (1932). Praca ta dotyczyła metod oceny wartości podstawowych stałych fizycznych na podstawie dostępnych pomiarów wykonywanych bardzo różnymi metodami wykorzystującymi różne zjawiska, w których opisie te stałe się pojawiają.

5 Wyznaczanie parametrów prostej: $y = ax + b$

Jako ostatni przykład zastosowania metody najmniejszych kwadratów przeanalizujemy pokrótce znajdowanie parametrów zależności $y = ax + b$ (parametrów prostej) gdy znane są wyniki pomiarów $\hat{y}_i \pm u_i$ dla *dokładnie znanych* wartości x_i – pokrótce, bo całość analizy jest analogiczna do analizy przeprowadzonej w przypadku średniej ważonej i współczynnika proporcjonalności. Dysponujemy wynikami pomiarów:

$$x_i \mapsto \hat{y}_i \pm u_i, i = 1, \dots, N$$

Zakładamy, że pomiary \hat{y}_i i \hat{y}_j są niezależne, gdy $i \neq j$, a między wielkościami y i x zachodzi związek $y = ax + b$. Wykorzystujemy metodę najmniejszych kwadratów do znalezienia najlepszej oceny wartości parametrów a i b , ich niepewności u_a i u_b oraz współczynnika kowariancji C_{ab} :

- Traktujemy każdy z wyników pomiarów \hat{y}_i jak zmienną losową o wariancji σ_i^2 .
- Poszukujemy minimum funkcji:

$$\mathcal{R}(a, b) = \sum_{i=1}^N \frac{(\hat{y}_i - ax_i - b)^2}{\sigma_i^2}$$

względem a i b .

Warunki konieczne minimum \mathcal{R} to znikanie obu pochodnych cząstkowych:

$$\frac{\partial \mathcal{R}}{\partial a} = 0, \quad \frac{\partial \mathcal{R}}{\partial b} = 0.$$

Warunki te prowadzą do układu równań spełnianych przez poszukiwane oceny \hat{a} i \hat{b} :

$$\begin{aligned} \hat{a} \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} + \hat{b} \sum_{i=1}^N \frac{x_i}{\sigma_i^2} &= \sum_{i=1}^N \frac{\hat{y}_i x_i}{\sigma_i^2} \\ \hat{a} \sum_{i=1}^N \frac{x_i}{\sigma_i^2} + \hat{b} \sum_{i=1}^N \frac{1}{\sigma_i^2} &= \sum_{i=1}^N \frac{\hat{y}_i}{\sigma_i^2}. \end{aligned}$$

Rozwiązaniem tego układu są:

$$\begin{aligned}
D &= \sum_{i=1}^N \frac{1}{\sigma_i^2} \sum_{j=1}^N \frac{x_j^2}{\sigma_j^2} - \left(\sum_{i=1}^N \frac{x_i}{\sigma_i^2} \right)^2 \\
\hat{a} &= \frac{1}{D} \left(\sum_{i=1}^N \frac{1}{\sigma_i^2} \sum_{j=1}^N \frac{\hat{y}_j x_j}{\sigma_j^2} - \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \sum_{j=1}^N \frac{\hat{y}_j}{\sigma_j^2} \right) \\
\hat{b} &= \frac{1}{D} \left(\sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} \sum_{j=1}^N \frac{\hat{y}_j}{\sigma_j^2} - \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \sum_{j=1}^N \frac{\hat{y}_j x_j}{\sigma_j^2} \right).
\end{aligned}$$

Otrzymane zmienne losowe \hat{a} i \hat{b} są liniowymi funkcjami zmiennych losowych \hat{y}_i , co prowadzi do następujących ocen ich wariancji i kowariancji:

$$\begin{aligned}
u_{\hat{a}}^2 &= \frac{1}{D} \sum_{i=1}^N \frac{1}{\sigma_i^2} \\
u_{\hat{b}}^2 &= \frac{1}{D} \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} \\
C_{\hat{a}\hat{b}} &:= \mathcal{E}((\hat{a} - a)(\hat{b} - b)) = -\frac{1}{D} \sum_{i=1}^N \frac{x_i}{\sigma_i^2}.
\end{aligned}$$

W ostatnim wzorze a i b oznaczają dokładne wartości obu parametrów.

Warto zwrócić uwagę na fakt, że wyznaczone współczynniki \hat{a} i \hat{b} nie są niezależnymi zmiennymi losowymi, bo na ogół $C_{\hat{a}\hat{b}} \neq 0$.

W §22 *slajdów do wykładu* podane są wzory “do obliczeń”, gdzie w miejsce σ_i podstawiono s_i – w ogólnym przypadku niepewności \hat{y}_i wyznaczanych **metodą B** zamiast s_i podstawiamy u_i . Podane są tam także wzory na oceny tych samych wielkości, gdy nie znamy niepewności \hat{y}_i , ale możemy przyjąć, że są one dla wszystkich \hat{y}_i takie same i ocenić ich wartość korzystając z przytoczonego wyżej twierdzenia – metodą najmniejszych kwadratów znalezione zostały dwa parametry \hat{a} i \hat{b} , a więc mianownik oceny s^2 (u^2) równy jest $N - 2$.

Rysunek 9 *slajdów do wykładu* ilustruje wpływ wartości u_i na obliczone parametry \hat{a} i \hat{b} dla tego samego zbioru wartości \hat{y}_i . Na ostatnim z tych przykładów – prawy dolny obrazek – wyznaczona wartość $a = -0,09 \pm 0,37$ jest zgodna z zerem. Jeżeli jednak, na tej podstawie, chcemy przyjąć, że y nie zależy od $x \rightarrow a = 0$, to za wartość \hat{y} nie możemy przyjąć obliczonej jednocześnie wartości $b = 4,7 \pm 1,3$. Po przyjęciu $a = 0$ za wartość \hat{y} przyjmujemy średnią ważoną wyników \hat{y}_i – tzn. musimy wykonać nowe obliczenia.

§22 *slajdów do wykładu* kończą uwagi dotyczące sposobu podejmowania decyzji, którą zmienną, x czy y , należy uznać za znaną *dokładnie*. Zwykle wartości obu są wynikami pomiarów, a więc tym samym są znane z niepewnościami. Kryterium jest proste i łatwe do zastosowania, jeśli wykona się wykres danych. Wtedy przybliżonej oceny parametru a można dokonać za pomocą linijki – tj. wykreślając przybliżony przebieg optymalnej prostej.

Wyznaczone parametry mogą służyć do wyznaczenia “poprawionych” wartości y_i w punktach pomiarowych lub obliczaniu wartości y w punktach x , dla których nie wykonywano pomiarów. O ile wyznaczanie wartości dla punktów x leżących wewnątrz badanego obszaru jest

bardzo “bezpieczne”, to należy bardzo ostrożnie traktować wyniki ekstrapolacji, czyli obliczeń dla x spoza obszaru, w którym wykonywano pomiary. Odpowiednie wartości y otrzymujemy po prostu podstawiając wyznaczone wartości \hat{a} i \hat{b} do wzoru $y = ax + b$. Ostatni punkt §22 *slajdów do wykładu* podaje także sposób wyznaczania niepewności tak otrzymanego wyniku – przytoczony tam wzór wynika z “prawa propagacji niepewności” (omówionego w §18 *slajdów do wykładu*).

Uwagi:

- Analogicznie do powyższych rozważań można metodą najmniejszych kwadratów znajdować także współczynniki zależności wielomianowej $y(x) = a_k x^k + \dots + a_1 x^1 + a_0$, gdy znane są wyniki pomiarów $x_i \mapsto \hat{y}_i \pm u_i; i = 1, \dots, N > k+1$. Oceny wartości parametrów $\hat{a}_j; j = 0, \dots, k$ otrzymuje się jako rozwiązania układu $k+1$ równań liniowych uzyskanych z przyrównania do zera pochodnych cząstkowych względem a_0, \dots, a_k funkcji:

$$\mathcal{R}(a_0, \dots, a_n) = \sum_{i=1}^N \frac{(\hat{y}_i - a_k x_i^k \dots - a_1 x_i - a_0)^2}{\sigma_i^2}$$

Wyznaczone w ten sposób $\hat{a}_0, \dots, \hat{a}_k$ są liniowymi funkcjami zmiennych losowych (wyników pomiarów) $\hat{y}_i; i = 1, \dots, N > k+1$.

- Wiele nieliniowych związków między wielkościami fizycznymi, po transformacji zmiennych można sprowadzić do postaci “ $y = ax + b$ ”, na przykład:

$$\begin{aligned} N(t) = N_0 \exp(-\lambda t) &\longrightarrow \ln(N) = -\lambda t + \ln(n_0) \\ y = Ax^B &\longrightarrow \ln(y) = B \ln(x) + \ln(A) \\ pV = C &\longrightarrow p = C \frac{1}{V} \\ T = 2\pi \sqrt{\frac{l}{g}} &\longrightarrow T^2 = \frac{4\pi^2}{g} l. \end{aligned}$$

Transformacja “prostująca” zależność pozwala zastosować wyprowadzone powyżej wzory do oceny parametrów na podstawie wyników pomiarów. Należy przy tym pamiętać o wyznaczeniu niepewności nowych, przetransformowanych zmiennych niezależnych (wyników pomiarów). Dla przykładu:

$$\begin{aligned} u(\ln(N)) &= \frac{u(N)}{N} \\ u\left(\frac{1}{V}\right) &= \frac{u(V)}{V^2} \\ u(T^2) &= 2Tu(T). \end{aligned}$$

Zadania do Wykładu 6

Zadanie 1. Neutron jest niestabilną cząstką elementarną, która po pewnym czasie ulega rozpadowi na proton, elektron i antyneutrino elektronowe. Wykonano 3 niezależne doświadczenia, w których mierzono czas życia neutronu i otrzymano następujące wyniki: (883 ± 30) s, (888 ± 3) s, (894 ± 5) s. Podaj najlepszą ocenę czasu życia neutronu i niepewność tej oceny.

Zadanie 2. Podczas zajęć pracowni fizycznej trzech studenci A, B i C wyznacжали gęstość trzech próbek metalu (każdy z nich badał jedną próbkę) dostarczonych przez asystenta. O asystencie wiadomo, że dysponuje tylko dwoma typami próbek i są to próbki żelaza i chromu. W tabelach można znaleźć, że gęstość żelaza wynosi $7,874 \text{ g/cm}^3$, a chromu $7,140 \text{ g/cm}^3$. Studenci otrzymali następujące rezultaty: $\rho_A = (7,095 \pm 0,150) \text{ g/cm}^3$, $\rho_B = (8,006 \pm 0,150) \text{ g/cm}^3$ i $\rho_C = (7,070 \pm 0,081) \text{ g/cm}^3$. Ponieważ przynajmniej dwóch studentów badało ten sam metal, znajdź najlepszą ocenę gęstości tego metalu i jej niepewność, wynikające z pomiarów wykonanych przez studentów.

Zadanie 3. Równanie stanu gazu doskonałego ma postać: $pV = nRT$, gdzie p oznacza ciśnienie, V objętość, n liczbę moli gazu, T temperaturę gazu w skali Kelvina, a R to uniwersalna stała gazowa. Student badał zmiany objętości V dla $n = 0,05$ mola gazu od jego temperatury T przy stałym ciśnieniu $p = 1015 \text{ hPa}$. Po ustaleniu temperatury T , pomiar objętości V powtarzał wielokrotnie i obliczał wartość średnią \bar{V} . Wyniki pomiarów zebrał w tabeli:

Temperatura T/K	293	298	303	308	313
Średnia objętość V/litr	1,196	1,207	1,243	1,274	1,293

Niepewności kolejnych wartości \bar{V} były bardzo zbliżone i wynosiły $u_v = 0,010$ litra. Przyjmij, że liczba moli n i ciśnienie p gazu znane są wystarczająco dokładnie. Wyznacz ocenę wartości stałej gazowej i niepewność tej oceny.

Zadanie 4. Studentka miała za zadanie wyznaczyć nieznaną opór R pewnego opornika. Do dyspozycji miała układ elektryczny, w którym mogła zmieniać napięcie na oporniku oraz amperomierz, którym dla ustalonego napięcia U mierzyła natężenie I prądu płynącego przez opornik. Wartość napięcia ustalana była z bardzo dużą dokładnością (przyjmij, że niepewność wartości U jest pomijalnie mała). Niepewność pomiaru natężenia prądu była niepomijalna i wynosiła u , zgodnie wartościami podanymi w tabeli. Zależność między natężeniem I i napięciem opisuje wzór: $U = RI$. Otrzymane wyniki studentka zapisała w tabeli:

Numer pomiaru	1	2	3	4	5
I [mA]	0,48	1,24	2,14	2,35	2,91
U [mV]	1,12	2,31	4,15	4,95	5,81
u [mA]	0,05	0,10	0,10	0,15	0,20

Wyznacz ocenę wartości oporu R oraz niepewność tej oceny.

Zadanie 5. Student badał przemianę izochoryczną gazu. Podgrzewał pewną ilość tego gazu w zamkniętym naczyniu i mierzył ciśnienie p i temperaturę t w skali Celsjusza i uzyskał wyniki podane w tabeli. Zakładając, że w warunkach doświadczenia gaz zachowuje się jak gaz doskonały, tzn. spełnia równanie:

$$\frac{pV}{T} = \text{const},$$

gdzie p jest jego ciśnieniem, V objętością, a T temperaturą w skali bezwzględnej, wyznacz ocenę temperatury zera bezwzględnego w skali Celsjusza i niepewność tej oceny. Zastosuj metodę najmniejszych kwadratów. Przyjmij, że temperatura była mierzona z niepewnością $u = 2^\circ\text{C}$, a ciśnienie było mierzone dostatecznie dokładnie.

Ciśnienie p (atmosfery)	1,0	1,4	1,6	2,0
Temperatura t ($^\circ\text{C}$)	36	158	223	350

Zadanie 6. Studentka przeprowadziła pomiar ciepła c parowania wody zdefiniowanego związkiem $Q = cm$, gdzie Q jest ilością energii, jaką należy dostarczyć masie m wody, żeby zamienić ją w parę. Ustawiła na wadze otwarty termos z pewną ilością wody i zanurzoną w wodzie grzałką o mocy $P = 600$ W. Po włączeniu grzałki, odczekała aż woda zacznie wrzeć i rozpoczęła pomiary masy m układu rejestrując wskazania wagi w równych odstępach czasu. Wyniki zestawiała w tabeli:

Czas t (minuty)	0	5	10	15	20
Wskazania wagi m (gramy)	1000	910	800	770	720

Przyjmij, że czas mierzony był wystarczająco dokładnie, także moc grzałki znana jest wystarczająco dokładnie i pozostawała stała podczas całego pomiaru. Wyznacz ciepło parowania wody c i niepewność uzyskanej oceny.