

# Uogólniona metoda najmniejszych kwadratów

## Generalized least squares method

Jacek Puchalski  
Główny Urząd Miar

Artykuł przedstawia uogólnione podejście dla dobrze znanej metody najmniejszych kwadratów stosowanej w praktyce metrologicznej. Wyznaczone niepewności punktów pomiarowych i korelacje między mierzonymi zmiennymi tworzą symetryczną macierz kowariancji, której odwrotność mnożona jest lewostronnie i prawostronnie przez wektory błędów obu zmiennych losowych i stanowi funkcję kryterialną celu. Aby uzyskać maksymalną wartość funkcji największej wiarygodności i rozwiązać złożony problemu minimalizacji funkcji kryterialnej, zaprezentowano oryginalny sposób wyznaczenia funkcji kryterialnej do postaci jednoargumentowej zależności obliczanej numerycznie, w której jedyną zmienną jest poszukiwany współczynnik kierunkowy prostej regresji. Artykuł zawiera podstawowe informacje o tego typu regresji liniowej, dla której najlepiej dopasowana prosta minimalizuje funkcję celu. Na przykładzie obliczeniowym pokazana jest pełna procedura dopasowania numerycznego prostej do danego zestawu punktów pomiarowych o zadanych niepewnościach i współczynnikach korelacji tworzących macierz kowariancji.

The paper presents a generalized approach for the well-known least squares method used in metrological practice. In order to solve the complex problem of minimizing the objective function to obtain the maximum value of the likelihood function, the original way of determining this function in the form of a unary relationship calculated numerically was presented. The article presents borderline cases with analytical solutions. The computational example shows the full procedure of numerical adjustment of a straight line to a given set of measurement points with given uncertainties and correlation coefficients forming the covariance matrix.

**Słowa kluczowe:** metoda najmniejszych kwadratów, regresja liniowa, niepewność, macierz kowariancji  
**Keywords:** least square method, linear regression, uncertainty, covariance matrix

## Wprowadzenie

Metoda najmniejszych kwadratów to metoda dopasowania na ogół modelu liniowego do danych pomiarowych, znana już od początku XIX wieku, a jej prekursorami byli francuski matematyk Adrien Marie Legendre oraz niemiecki matematyk i astronom Carl Friedrich Gauss. Początkowo stosowano ją w geodezji i astronomii, a obecnie we wszystkich pomiarach fizycznych, które wymagają zastosowania szeroko pojętego modelu liniowego. W tych rozważaniach skoncentrujemy się wyłącznie na dopasowaniu liniowej charakterystyki, opisanej linią prostą w kartezjańskim układzie współrzędnych:  $y = ax + b$ , gdzie  $x$  jest zmienną niezależną,  $a$  jest współczynnikiem kierunkowym prostej oraz  $b$  jest wyrazem wolnym, co ma najczęściej miejsce w praktyce metrologicznej [1, 2], gdy dokonujemy pomiarów w wybranych punktach charakterystyki parametru fizycznego, np. rezystancji elementu oporowego w funkcji zmiennej niezależnej jaką jest np. temperatura otoczenia, którą regulujemy w określonym zakresie. A zatem jest to regresja prosta, tj.  $y$  względem  $x$ , gdzie  $x$  jest zmienną niezależną,

w której, stosując metodę najmniejszych kwadratów, wyznaczamy parametry prostej najlepiej dopasowanej do punktów pomiarowych. Rezultatem tego dopasowania jest wyznaczenie wartości  $y$  w dowolnym punkcie charakterystyki z niepewnością rozszerzoną:

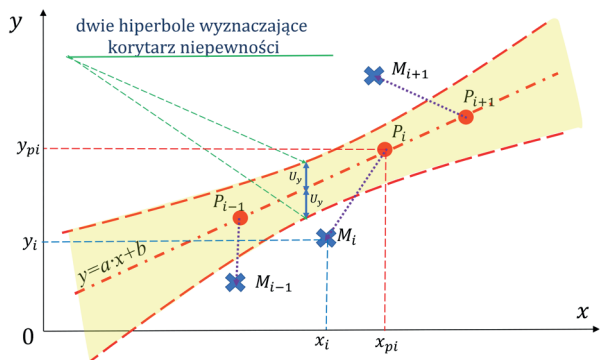
$$y = ax + b \pm U_y \quad (1)$$

Ze względu na fakt, że podczas wielokrotnych pomiarów współrzędne punktów są obciążone błędami powodującymi ich „rozmycie”, to oszacowane standardowe niepewności pomiarów obu zmiennych losowych reprezentowanych przez współrzędne  $x_i$  i  $y_i$ , z odpowiednimi niepewnościami  $u(x_i)$  i  $u(y_i)$ , skutkują tym, że precyzyjnie wyznaczone parametry prostej też ulegają „rozmyciu” odzwierciedlone w standardowych niepewnościach  $u_a$  i  $u_b$  oraz, w ogólności, we współczynniku korelacji  $\rho_{ab}$  między zmiennymi losowymi  $a$  i  $b$ . Niepewność rozszerzona dla prostej regresji wyraża się poprzez sumę skorelowanych rozkładów dla parametrów prostej  $a$  (rozkład dla  $a$  jest mnożony przez określoną wartość  $x$ ) i  $b$  mnożoną przez

współczynnik rozszerzenia wynikający z rozkładu  $t$ -Studenta dla prawdopodobieństwa  $p = 0,95$  ( $\alpha = 1 - p = 0,05$ ) i  $n - 2$  stopniami swobody:

$$U_y = t_{1-\alpha/2, n-2} \sqrt{x^2 u_a^2 + 2\rho_{ab} x u_a u_b + u_b^2} \quad (2)$$

Nietrudno jest wykazać, że „korytarz” niepewności jest ograniczony dwoma hiperbolami, tak jak na rys. 1.



■ Rys. 1. Punkty pomiarowe i ich projekcje

Jak widać z powyższego każdemu punktowi pomiarowemu  $M_i$  o parze współrzędnych  $(x_i, y_i)$  odpowiada punkt  $P_i$  o parze współrzędnych  $(x_{pi}, y_{pi})$ , który powstaje w wyniku rzutowania, pod odpowiednio dobranym kątem, punktu pomiarowego na prostą regresji  $y_{pi} = ax_{pi} + b$ , po to aby wyznaczyć błędy związane z jej dopasowaniem, tj.  $\Delta x_i = x_i - x_{pi}$  oraz  $\Delta y_i = y_i - y_{pi}$ . W przyjętym modelu założymy również, w ogólności, że powstałe w ten sposób błędy są skorelowane i to nie tylko w obrębie jednej współrzędnej  $x$  czy też  $y$ , ale skorelowane są również między sobą błędy w obu współrzędnych. Zastosowana tu metoda o akronimie WTLS (Weighted Total Least Square) jest uogólnieniem metody WLS (Weighted Least Square), tzw. ważonej metody najmniejszych kwadratów, gdzie minimalizowana jest suma kwadratów błędów normalizowanych do kwadratów niepewności odpowiednio dla obu współrzędnych (wagi są równe odwrotności kwadratów niepewności). Metoda WTLS, w przeciwieństwie do WLS, wymaga zastosowania rachunku wektorowego z zakresu algebry liniowej i uwzględnienia macierzy kowariancji występujących tu korelacji, podczas gdy w metodzie WLS, gdzie nie uwzględnia się korelacji współrzędnych, macierz kowariancji redukuje się do wyrazów na przekątnej, wystarcza zastosowanie rachunku algebraicznego w postaci sum. Przy okazji należy wspomnieć, że najbardziej znaną metodą najmniejszych kwadratów jest szczególnie przypadek metody WLS o akronimie OLS (Ordinary Least Square), w której zakłada się, że wszystkie niepewności są identyczne (wagi są równe jeden), co prowadzi bezpośrednio do minimalizacji sumy kwadratów odległości od prostej regresji [1, 2].

## Efektywna macierz kowariancji $U_{yeff}$

W tej części zostanie przedstawiony sposób zastąpienia macierzy kowariancji uwzględniającej losowy charakter obu współrzędnych kartezjańskich i ich korelację przez zastosowanie efektywnej macierzy kowariancji dla regresji prostej  $y$  względem zmiennej niezależnej  $x$  [3]. W tym celu utworzymy zmienną losową wektorową o rozmiarze  $2n$  obejmującą wszystkie współrzędne punktów pomiarowych  $Z = [X^T, Y^T]^T$ , gdzie wektory współrzędnych punktów pomiarowych są określone przez  $X = [x_1, \dots, x_i, \dots, x_n]^T$  i  $Y = [y_1, \dots, y_i, \dots, y_n]^T$  są rozmiaru  $n$  oraz wartości współrzędnych na prostej regresji spełniają warunek  $Y_p = aX_p + bQ = aX_p + b$ , gdzie oznaczono jednostkowy wektor  $Q = Q_{n \times 1} = [1, \dots, 1]^T$  rozmiaru  $n$ , które tworzą  $2n$ -wymiarową wektorową wartość oczekiwaną  $Z_p = [X_p^T, Y_p^T]^T$  wielowymiarowego rozkładu Gaussa, stanowiącego funkcję największej wiarygodności:

$$f(\Delta Z) = \frac{1}{(2\pi)^n \sqrt{\det(U_Z)}} \exp\left\{-\frac{1}{2} \phi(\Delta Z)\right\} \quad (3)$$

gdzie funkcja kryterialna  $\phi(\Delta Z) = \Delta Z^T U_Z^{-1} \Delta Z$ , a wektor  $\Delta Z = Z - Z_p$  jest łącznym wektorem błędów zawierającym błędy obu współrzędnych, zaś symetryczna macierz kowariancji  $U_Z$  o rozmiarze  $2n \times 2n$  jest zdefiniowana następująco ( $\det(U_Z) > 0$ ):

$$U_Z = \begin{bmatrix} u^2(z_1) & \dots & \rho_{1,2n} u(z_1) u(z_{2n}) \\ \dots & \cdot & \dots \\ \rho_{1,2n} u(z_{2n}) u(z_1) & \dots & u^2(z_{2n}) \end{bmatrix} \quad (4)$$

Oczywiście  $u(z_i) = u(x_i)$  dla  $0 \leq i \leq n$  oraz  $u(z_i) = u(y_i)$  dla  $n+1 \leq i \leq 2n$ , a symbolem  $-1 \leq \rho_{kl} \leq 1$  oznaczono współczynniki korelacji Pearsona korelujących zmiennych losowych  $z_k$  i  $z_l$ ,  $1 \leq k, l \leq 2n$ . W postaci skonsolidowanej, uwzględniającej wydzielone bloki symetryczne, macierze:  $U_Z$  i jej odwrotność  $U_Z^{-1}$  mogą być zapisane:

$$U_Z = \begin{bmatrix} U_X & U_{XY} \\ U_{XY}^T & U_Y \end{bmatrix} \quad (5)$$

$$U_Z^{-1} = \begin{bmatrix} V_1 & V_3 \\ V_3^T & V_2 \end{bmatrix} \quad (6)$$

W ogólności macierze  $U_X, U_Y (V_1, V_2)$  o rozmiarze  $n \times n$  są symetryczne – powinny być pozytywnie określone, natomiast korelacje krzyżowe zawarte w macierzy w  $U_{XY}$  mogą prowadzić do warunku  $U_{XY} \neq U_{XY}^T$  (analogicznie  $V_3 \neq V_3^T$ ), co oznacza, że te wyszczególnione wyżej bloki macierzowe nie zawsze są symetryczne.

Uzyskanie warunku maksymalizacji funkcji największej wiarygodności  $f(\Delta Z)$ , która jest wielowymiarowym rozkładem Gaussa prowadzi do minimalizacji wykładnika potęgi tej funkcji, tj. funkcji kryterialnej  $\phi(\Delta Z) = \Delta Z^T U_Z^{-1} \Delta Z$  oraz zapisanego wektorami równania prostej regresji, spełniającej rolę wartości oczekiwanej:

$$\begin{cases} \phi(\Delta Z) \rightarrow \min \\ Y_p = aX_p + b \end{cases} \quad (7)$$

Po zastąpieniu drugiego z równań przez wprowadzenie wektorów błędów  $\Delta X$  i  $\Delta Y$ , tj.  $Y - Y_p - Y = aX - aX_p - b - aX$ , a stąd  $\Delta Y = a\Delta X + E$ , gdzie oznaczono  $E = Y - aX - b$  i po wstawieniu tej zależności do zależności na  $\phi(\Delta Z)$ , pamiętając że  $\Delta Z = [\Delta X^T, \Delta Y^T]^T$ , otrzymuje się następującą zależność:

$$\phi(\Delta X, a, b) = \Delta X^T V \Delta X + 2E^T (V_3^T + aV_2) \Delta X + E^T V_2 E \quad (8)$$

if symetryczna macierz  $V$  jest zdefiniowana  $V = V_1 + a(V_3^T + V_3) + a^2 V_2$ . Okazuje się, że dla nieskorelowanych wielkości, gdy macierz  $U_Z$  jest diagonalna, macierz  $V$  jest również diagonalna z elementami dodatnimi na przekątnej. W rezultacie funkcja kryterialna  $\phi(\Delta X, a, b)$  jest sumą form kwadratowych i ta hiperparaboliczna zależność osiąga minimum w wierzchołku hiperparaboli, tj. we wszystkich punktach dla których poszczególne formy kwadratowe osiągają minima. Analogiczny dowód dla niediagonalnej macierzy  $V$ , którą diagonalizujemy, można odnaleźć w [3]. W rezultacie, jeśli macierz  $V$  jest pozytywnie określona, tzn. wszystkie wartości własne są dodatnie, to minimum znajdujące się w wierzchołku hiperparaboli występuje dla warunku zerowego  $n$ -wymiarowego gradientu

$\nabla \phi(\Delta X, a, b) = \mathbf{0}_{n \times 1}$ , tj. dla  $\Delta X^T = -E^T (V_3^T + aV_2) V^{-1}$ . Spełniona jest zatem relacja:

$$\phi(\Delta X, a, b) \geq \phi(a, b) = E^T [V_3 - (V_3^T + aV_2) V^{-1} (V_3 + aV_2)] E \quad (9)$$

Funkcja kryterialna  $\phi(a, b)$  może być zinterpretowana jako funkcja wyznaczona dla regresji prostej  $y$  względem  $x$ , przy czym występuje w niej efektywna odwrotna do macierzy kowariancji macierz symetryczna, zależna od współczynnika kierunkowego prostej regresji, określona jako:

$$U_{\text{eff}}^{-1} = V_2 - (V_3^T + aV_2) V^{-1} (V_3 + aV_2) \quad (10)$$

W szczególnym przypadku, gdy macierz  $V_3$  jest symetryczna, tj.  $V_3 = V_3^T$ , wówczas  $U_{\text{eff}}^{-1} = (V_1 V_2 - V_3 V_3) V^{-1}$  oraz gdy  $V_3 = \mathbf{0}_{n \times n}$  (macierz zerowa o rozmiarze  $n \times n$ ), wtedy  $U_{\text{eff}}^{-1} = a^2 U_X + U_Y$ . W rezultacie dla nieskorelowanych wielkości, kiedy macierze  $U_X$  i  $U_Y$  są diagonalne,  $U_{\text{eff}}$  posiada wyłącznie niezerowe wartości na przekątnej równe:  $[U_{\text{eff}}^{-1}]_{ii} = a^2 u^2(x_i) + u^2(y_i)$ .

## Konsekwencje wyboru punktów na prostej regresji

Minimalizacja związana z wyborem punktów na prostej, który minimalizuje funkcję kryterialną, prowadzi do liniowej zależności między błędami w kierunku  $OX$   $\Delta X = -V^{-1} (V_3 + aV_2) E$  i błędami w kierunku  $OY$   $\Delta Y = -V^{-1} (V_1 + aV_3^T) E$ , poprzez macierz współczynników wagowych:

$$\beta = -(V_1 + aV_3^T) (V_3 + aV_2)^{-1} \quad (11)$$

lub

$$\beta^{-1} = -(V_1 + aV_3^T)^{-1} (V_3 + aV_2)$$

Wówczas prawdziwe są zależności:  $\Delta Y = \beta \Delta X$  i  $\Delta X = \beta^{-1} \Delta Y$ .

Odpowiednio rozwiązanie dla regresji prostej  $y$ , względem zmiennej niezależnej  $x$ , otrzymujemy gdy  $\beta^{-1} \rightarrow \mathbf{0}_{n \times n}$  ( $U_X \rightarrow \mathbf{0}_{n \times n}$ , np. zakładając bliskie zera wartości na przekątnej tej macierzy, co oznacza, że wartości odwrotne na przekątnej dążą do nieskończoności), natomiast rozwiązanie dla regresji odwrotnej  $x$ , względem zmiennej niezależnej  $y$ , otrzymujemy gdy  $\beta^{-1} \rightarrow \mathbf{0}_{n \times n}$  ( $U_Y \rightarrow \mathbf{0}_{n \times n}$ , np. dotyczy to bliskich zera wartości na przekątnej tej macierzy, powodując że wartości odwrotne dążą do nieskończoności).

Należy zauważyć, że dla całkowicie nieskorelowanych danych macierz  $\beta$  zawiera wyłącznie wartości niezerowe na przekątnej, które wynoszą:  $-a^{-1} u^2(y_i) u^{-2}(x_i)$  i stanowią współczynniki kierunkowe prostych przechodzących przez punkty pomiarowe i przecinających prostą regresji w punktach względem których obliczane są poszczególne błędy będące składowymi wektorów  $\Delta X$  i  $\Delta Y$ .

## Wzór na jednoargumentową funkcję kryterialną

Wprowadzając oznaczenia w postaci parametrów, tj. wielkości skalarnych otrzymanych przez obustronne mnożenie macierzy wektorami:

$$\begin{aligned} S &= Q^T U_{\text{eff}}^{-1} Q = \sum_{i=1}^n \sum_{j=1}^n [u_{\text{eff}}^{-1}]_{ij} \\ S_X &= X^T U_{\text{eff}}^{-1} Q = Q^T U_{\text{eff}}^{-1} X \\ S_Y &= Y^T U_{\text{eff}}^{-1} Q = Q^T U_{\text{eff}}^{-1} Y \\ S_{XX} &= X^T U_{\text{eff}}^{-1} X \\ S_{YY} &= Y^T U_{\text{eff}}^{-1} Y \\ S_{XY} &= X^T U_{\text{eff}}^{-1} Y = Y^T U_{\text{eff}}^{-1} X \end{aligned} \quad (12a-f)$$

gdzie symbolem  $[u_{\text{eff}}^{-1}]_{ij}$  oznaczono elementy macierzy  $U_{\text{eff}}^{-1}$ , możemy funkcję kryterialną przedstawić w postaci formy kwadratowej względem współczynnika kierunkowego i wyrazu wolnego prostej regresji:

$$\begin{aligned} \phi(a,b) &= a^2 S_{XX} + 2abS_X - 2aS_{XY} \\ &+ b^2 S - 2bS_Y + S_{YY} \end{aligned} \quad (13)$$

Przy założeniu, że parametr skalarny  $S > 0$ , zależność kwadratowa od  $b$  prowadzi do lokalnego minimum przy warunku:

$$\frac{\partial \phi(a,b)}{\partial b} = 0 \quad (14)$$

W związku z tym powyższy warunek prowadzi do zależności między  $a$  i  $b$ :

$$b = (S_Y - aS_X) / S \quad (15)$$

W rezultacie, po podstawieniu, otrzymujemy funkcję zależną od jednej zmiennej  $a$ , gdyż parametry z (12a-f) zależą, poprzez  $\mathbf{U}_{\text{Yeff}}^{-1}$ , wyłącznie od współczynnika kierunkowego prostej regresji:

$$\begin{aligned} \phi(a) &= a^2(S_{XX} - S_X^2/S) + 2a(S_X S_Y / S - S_{XY}) + \\ &+ S_{YY} - S_Y^2/S \end{aligned} \quad (16)$$

która spełnia warunek:

$$\phi(\Delta \mathbf{X}, a, b) \geq \phi(a, b) \geq \phi(a) \quad (17)$$

W związku z powyższym, aby odnaleźć maksimum funkcji największej wiarygodności wystarczy odnaleźć minimum funkcji  $\phi(a)$ , co analitycznie jest możliwe, ale tylko w szczególnych przypadkach. W ogólności należy w danym przedziale przewidywanych wartości współczynnika kierunkowego znaleźć numerycznie wartości tej funkcji, generując serie punktów z krokiem np.  $\Delta a$ , a następnie wykreślić charakterystykę tej funkcji. Wykorzystać do tego można skoroszyt z arkuszami EXCEL, środowisko programistyczne MATLAB czy R albo inne dedykowane.

## Algorytm numeryczny

Wyznaczenie charakterystyki  $\phi(a)$ , która to funkcja ma uwidocznić minimum dla poszukiwanego  $a$ , wymaga wyznaczenia jej wartości w danym przedziale zmienności  $[a_L, a_H]$ , np. przez zastosowanie inkrementacji od przyjętej dolnej wartości przedziału  $a_L$  z odpowiednio dobranym krokiem  $\Delta a$ :  $a = a_L + i\Delta a$ , gdzie liczba  $i$  kroków zawarta jest w przedziale:  $0 \leq i \leq [(a_H - a_L)/\Delta a]$ , aby osiągnąć górną wartość przedziału  $a_H$ , próbując cały zadany przedział. Z kolei korzystając z (15) możemy wyznaczyć analogiczną charakterystykę  $\phi(b)$ .

Końcowym efektem numerycznych obliczeń jest uzyskanie charakterystyk  $\phi(a)$  i  $\phi(b)$ , z których odczytujemy wartości minimalizujące funkcję  $\phi(\Delta \mathbf{X}, a, b)$ , przez relację (17) w punktach  $a_N$  i  $b_N$  z błędami numerycznymi określonymi w [3].

Dodatkowo z prawa propagacji niepewności otrzymujemy zależność:

$$\begin{aligned} \mathbf{U}_{ab} &= \mathbf{C} \mathbf{U}_Z \mathbf{C}^T \\ \text{i } \mathbf{U}_{ab} &= \begin{bmatrix} u_a^2 & \rho_{ab} u_a u_b \\ \rho_{ab} u_a u_b & u_b^2 \end{bmatrix} \end{aligned} \quad (18a-b)$$

gdzie macierz wrażliwości  $\mathbf{C}$  o rozmiarze  $2n \times 2n$  jest określona przez:

$$\mathbf{C} = \begin{bmatrix} \frac{\partial a}{\partial \mathbf{Z}}, \frac{\partial b}{\partial \mathbf{Z}} \end{bmatrix}^T = \begin{bmatrix} \frac{\partial a}{\partial x_1} \dots \frac{\partial a}{\partial x_n}, \frac{\partial a}{\partial y_1} \dots \frac{\partial a}{\partial y_n} \\ \frac{\partial b}{\partial x_1} \dots \frac{\partial b}{\partial x_n}, \frac{\partial b}{\partial y_1} \dots \frac{\partial b}{\partial y_n} \end{bmatrix} \quad (19)$$

Wyznaczenie macierzy  $\mathbf{C}$ , tj. pochodnych cząstkowych, zrealizowane jest przez numeryczne różniczkowanie przy wykorzystaniu formuł:

$$\frac{\partial c}{\partial z_i} \cong \frac{c_N(z_i + h) - c_N(z_i - h)}{2h} \quad (20)$$

gdzie  $c = a$  albo  $c = b$ , które wymagają wyznaczenia  $4 \times n$  rozwiązań dla współczynnika kierunkowego  $a$  i wyrazu wolnego  $b$  po modyfikacji współrzędnych o niewielki przyrost  $h$ , tj.  $x_i \pm h$  i  $y_i \pm h$ .

## Graniczne przykłady regresji dwóch współrzędnych obciążonych błędami pomiarów

W sytuacji gdy jedna z macierzy  $\mathbf{U}_X$  lub  $\mathbf{U}_Y$  posiada bliskie zeru wartości na przekątnej, tj. kwadraty niepewności, przy założeniu, że macierz  $\mathbf{U}_{XY} = \mathbf{0}_{n \times n}$ , tzn. brak jest korelacji krzyżowych między zmiennymi losowymi reprezentowanymi przez obie współrzędne  $x_i$  i  $y_i$ , mamy do czynienia z granicznym przejściem do regresji prostej albo odwrotnej, w której tylko jedna ze współrzędnych jest wyznaczana z pomiarów.

## Rozwiązanie dla prostej regresji y względem zmiennej niezależnej x

W tym przypadku zmienna  $x$  jest zmienną niezależną, a  $y$  jest zmienną zależną i jej wartości są mierzone z błędami pomiaru odzwierciedlone w niepewnościach i współczynnikach korelacji opisanych przez macierz kowariancji  $\mathbf{U}_Y$ . Wtedy w granicy przy  $\mathbf{U}_X \rightarrow \mathbf{0}_{n \times n}$  macierz efektywna  $\mathbf{U}_{\text{Yeff}}^{-1} \rightarrow \mathbf{U}_Y^{-1}$ , a macierz współczynników wagowych  $\boldsymbol{\beta}^{-1} \rightarrow \mathbf{0}_{n \times n}$  i funkcja kryterialna przyjmuje postać:

$$\phi(a,b) = \mathbf{E}^T \mathbf{U}_Y^{-1} \mathbf{E} \quad (21)$$

Należy zauważyć, że wektor błędów  $\mathbf{E}$  z pomiarów zawiera wartości błędów w kierunku  $0\mathbf{Y}$ , tj.  $\mathbf{E} = \Delta\mathbf{Y}$ . Korzystając z oznaczeń z (12a-f), po zamianie efektywnej macierzy kowariancji na macierz  $\mathbf{U}_Y$ , otrzymuje się z (13) rozwiązanie w postaci wyznaczników na wierzchołek dwu-parametrowej paraboli zależnej od  $\mathbf{a}$  i  $\mathbf{b}$ :

$$\mathbf{a} = \frac{\Delta_a}{\Delta}, \quad \mathbf{b} = \frac{\Delta_b}{\Delta} \quad (22a-b)$$

gdzie:  $\Delta_a = SS_{XY} - S_X S_Y$ ,  $\Delta_b = S_Y S_{XX} - S_X S_{XY}$  i  $\Delta = SS_{XX} - (S_X)^2$  (zakładamy  $\Delta > 0$ ).

Z kolei parametry macierzy kowariancji  $\mathbf{U}_{ab}$  wyrażają się poprzez wzory:

$$\begin{aligned} u_a^2 &= \frac{S}{\Delta}, \quad u_b^2 = \frac{S_{XX}}{\Delta} \\ \rho_{ab} u_a u_b &= -\frac{S_X}{\Delta} \quad \text{i} \quad \rho_{ab} = -\frac{S_X}{\sqrt{SS_{XX}}} \end{aligned} \quad (23a-d)$$

## Rozwiązanie dla regresji odwrotnej $x$ względem zmiennej niezależnej $y$

W tym przypadku zmienna  $y$  jest zmienną niezależną, a  $x$  jest zmienną zależną i jej wartości są mierzone z błędami pomiaru odzwierciedlone w niepewnościach i współczynnikach korelacji opisanych przez macierz kowariancji  $\mathbf{U}_X$ . Wtedy w granicy przy  $\mathbf{U}_Y \rightarrow \mathbf{0}_{n \times n}$  macierz efektywna  $\mathbf{U}_{Y\text{eff}}^{-1} \rightarrow \mathbf{U}_X^{-1} \mathbf{a}^{-2}$ , a macierz współczynników wagowych  $\boldsymbol{\beta} \rightarrow \mathbf{0}_{n \times n}$  i funkcja kryterialna przyjmuje postać:

$$\phi(\mathbf{a}, \mathbf{b}) = \mathbf{E}^T \mathbf{U}_X^{-1} \mathbf{a}^{-2} \mathbf{E} = (-\mathbf{a}^{-1} \mathbf{E}^T) \mathbf{U}_X^{-1} (-\mathbf{a}^{-1} \mathbf{E}) \quad (24)$$

Minimalizacja (24) w tym przypadku jest wykonywana ze względu na błędy pomiarów występujące wyłącznie w kierunku  $0\mathbf{X}$ , tj. błędy  $\Delta\mathbf{X} = \mathbf{X} + \mathbf{a}^{-1}\mathbf{b} - \mathbf{a}^{-1}\mathbf{Y} = -\mathbf{E}\mathbf{a}^{-1}$ . A zatem z punktu widzenia oznaczeń i rozwiązania problemu minimalizacji (21) i (24) dokonujemy minimalizacji analogicznej do regresji  $y$  względem niezależnej zmiennej  $x$ , opisanej wyżej poprzez zamianę oznaczeń  $\mathbf{Y}$  na  $\mathbf{X}$  oraz macierzy  $\mathbf{U}_Y$  na  $\mathbf{U}_X$ . W efekcie otrzymujemy rozwiązanie, które opisywane jest prostą  $\mathbf{X}_p = \mathbf{a}_x \mathbf{Y}_p + \mathbf{b}_x$  w postaci analogicznej do (22):

$$\mathbf{a}_x = \frac{\Delta_{ax}}{\Delta_x}, \quad \mathbf{b}_x = \frac{\Delta_{bx}}{\Delta_x} \quad (25a-b)$$

gdzie zastosowano nowe oznaczenia:  $\Delta_{ax} = WW_{XY} - W_X W_Y$ ,  $\Delta_{bx} = W_X W_{YY} - W_Y W_{XY}$  i  $\Delta_x = WW_{YY} - (W_Y)^2$  (zakładamy  $\Delta_x > 0$ ).

Zwróćmy uwagę, że niepewności charakteryzują wyłącznie zmienną  $\mathbf{X}$ . Stąd parametry macierzy kowariancji  $\mathbf{U}_{axbx}$  są analogicznie do (23a-d):

$$\begin{aligned} u_{ax}^2 &= \frac{W}{\Delta_x}, \quad u_{bx}^2 = \frac{W_{YY}}{\Delta_x} \\ \rho_{abx} u_{ax} u_{bx} &= -\frac{W_Y}{\Delta_x} \quad \text{i} \quad \rho_{abx} = -\frac{W_Y}{\sqrt{WW_{YY}}} \end{aligned} \quad (26a-d)$$

W tym momencie należy zauważyć, że w naszym przykładzie pomiary wykonujemy mierząc wartości rezystancji w funkcji temperatury (regresja  $y$  – *mierzona rezystancja* względem niezależnej zmiennej  $x$  – *mierzona temperatura*), to temperatura jest zmienną niezależną. Zastosowane poniżej matematyczne rozwiązanie dotyczy sytuacji, gdy w trakcie pomiarów niepewności temperatury są dominujące w porównaniu do niepewności rezystancji i dlatego też stosuje się dla regresji  $x$  względem zmiennej niezależnej  $y$  powyższe, z punktu widzenia algebry łatwiejsze, rozwiązanie. Aby stwierdzić, jak dominujące niepewności temperatury wpływają na korytarz niepewności mierzonej rezystancji, należy zastosować macierzowe prawo propagacji niepewności korzystając z zależności:

$$\mathbf{U}_{ab} = \begin{bmatrix} \frac{\partial a}{\partial a_x} & \frac{\partial a}{\partial b_x} \\ \frac{\partial b}{\partial a_x} & \frac{\partial b}{\partial b_x} \end{bmatrix} \begin{bmatrix} u_{ax}^2 & \rho_{abx} u_{ax} u_{bx} \\ \rho_{abx} u_{ax} u_{bx} & u_{bx}^2 \end{bmatrix} \begin{bmatrix} \frac{\partial a}{\partial a_x} & \frac{\partial b}{\partial a_x} \\ \frac{\partial a}{\partial b_x} & \frac{\partial b}{\partial b_x} \end{bmatrix} \quad (27)$$

gdzie:  $\mathbf{a} = \mathbf{a}_x^{-1} = \Delta_x / \Delta_{ax}$  i  $\mathbf{b} = -\mathbf{b}_x \mathbf{a}_x^{-1} = -\Delta_{bx} / \Delta_{ax}$ , co po uproszczeniu prowadzi do następujących formuł:

$$\begin{aligned} u_a &= \mathbf{a}^2 u_{ax} = \mathbf{a}^2 \sqrt{\frac{W}{\Delta_x}} \\ u_b &= |\mathbf{a}| \sqrt{\frac{b^2 W + W_{YY} - 2W_Y b}{\Delta_x}} \\ \rho_{ab} u_a u_b &= \mathbf{a}^3 \frac{bW - W_Y}{\Delta_x} \end{aligned} \quad (28a-c)$$

## Przykład obliczeniowy

Symulacje wykonano na przykładzie wzorcowania pirometru radiacyjnego, dla którego zmierzono charakterystykę w pięciu punktach pomiarowych. Wzorcowe wartości mierzone reprezentuje współrzędna  $x$  wraz z niepewnościami  $\mathbf{u}(x)$ , a mierzone przez wzorcowany pirometr współrzędna  $y$  i niepewności  $\mathbf{u}(y)$  (wyniki zawiera tab. 1). Dodatkowo przyjęto, że pomiary są skorelowane, tzn. istnieje niewielka, słaba korelacja  $\rho_{xy} = 0,2$ . Należy zauważyć że w tym przypadku macierze kowariancji są diagonalne, a próba rozwiązania analitycznego tego problemu nie powiedzie się, gdyż rozwiązanie dla skorelowanych zmiennych wymaga zastosowania numerycznej metody iteracyjnej Yorka [1]. Implementacji ww. proponowanego algorytmu numerycznego dokonano w środowisku MATLAB, choć równie dobrze może być to MATHCAD, STATISTICA, R czy nawet EXCEL.

Tab. 1. Współrzędne punktów pomiarowych wraz z niepewnościami z opcjonalną słabą korelacją

Numer punktu pomiarowego	$x_i$ °C	$u(x_i)$ °C	$y_i$ °C	$u(y_i)$ °C	$\rho_{xy}$
1	30	0,577	29,8	0,115	0,2
2	50	0,577	51,4	0,808	
3	70	0,577	70,3	0,173	
4	90	0,577	89,1	0,520	
5	110	0,577	108,4	0,924	

Zastosowany skrypt w MATLAB odczytuje dane wielkości wejściowych z arkusza EXCEL (tab. 1), a następnie wyznacza charakterystykę funkcji  $\phi(a)$ , opcjonalnie  $\phi(b)$  oraz umożliwia określenie minimum w środowisku MATLAB, np. użycie funkcji Zoom In. Na rys. 2 wykreślono dwie charakterystyki bez uwzględnienia korelacji ze wskazaniem minimum wartości funkcji i odpowiadających im parametrów  $a$  i  $b$ .

Następnie drugi skrypt wyznacza z (18) macierz kowariancji  $U_{ab}$ , używając numerycznego różniczkowania (20) w celu wyznaczenia współczynników wrażliwości (19). Analogicznych obliczeń wykonano z uwzględnieniem korelacji.

Wyniki obliczeń podano w poniższych tabelkach, ograniczając liczbę poprawnych cyfr znaczących z uwzględnieniem oszacowanych błędów numerycznych.

Wyniki w tab. 2 uwidaczniają nieznaczne zmiany w parametrach prostej, a także zmniejszenie parametrów niepewności współczynnika kierunkowego i wyrazu wolnego oraz zwiększenia wartości bezwzględnej współczynnika korelacji. Występuje tu silniejsza ujemna korelacja, dla uwzględnienia dodatniej korelacji mierzonych

współrzędnych. Oznacza to, że korytarz niepewności będzie węższy dla wartości mierzonych z uwzględnieniem korelacji.

W tab. 3 zaprezentowano porównanie wartości parametrów prostej wyznaczonych numerycznie z wartościami obliczonymi z analitycznych formuł. Zerowe wartości niepewności dla macierzy  $U_x$  i  $U_y$  zostały zasymulowane przez wstawienie dla zerowych wartości niepewności  $u^2(x_i) = 10^{-10}$  lub  $u^2(y_i) = 10^{-10}$ , aby można było odwrócić macierz  $U_z$ .

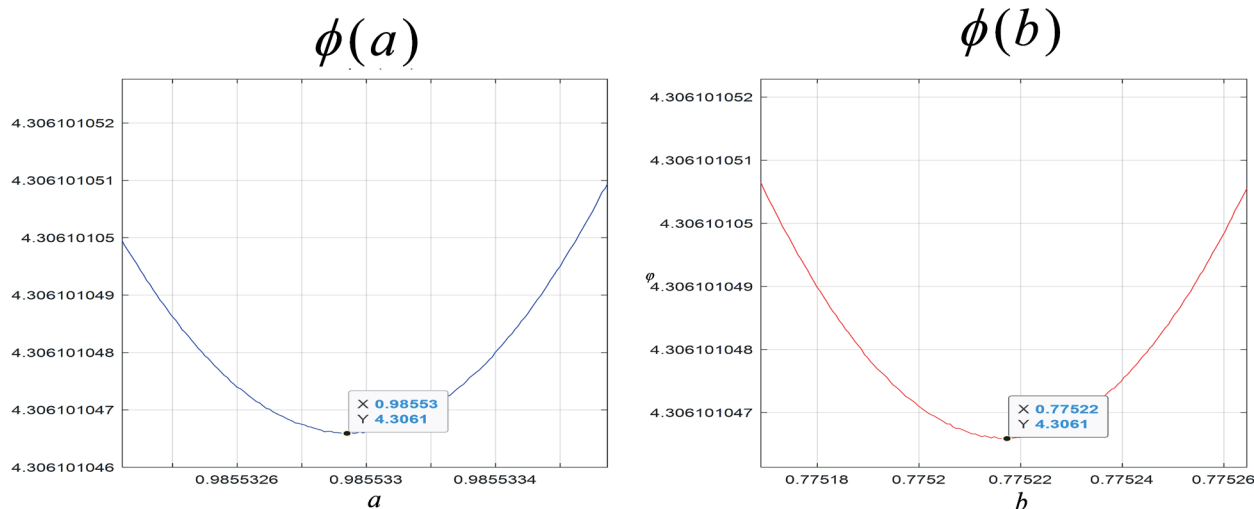
Z tab. 3 wynika, że uzyskano zbieżność rozwiązań zgodnymi z prawie czterema cyframi dziesiętnymi. Z kolei w tab. 4 ujęto wartości zmierzone przez wzorcowany pirometr z niepewnościami rozszerzonymi.

Jak pokazano w tab. 4 korytarz niepewności jest nawet do 20 % szerszy dla pomiarów bez uwzględnienia korelacji, w porównaniu do pomiarów uwzględniających płytką korelacją.

W tab. 4 zawarte są również wyniki niepewności rozszerzonej dla uwzględnienia niepewności tylko jednej zmiennej  $x$  albo  $y$ . Przy uwzględnieniu wyłącznie niepewności zmiennej  $y$ , przy bliskich zera niepewnościach wzorca otrzymano najwęższy korytarz niepewności.

## Podsumowanie

Przedstawiono nowe podejście do regresji liniowej reprezentującej optymalną prostą regresji w ogólnym przypadku, gdy obie zmienne (losowe), reprezentowane przez obie współrzędne kartezjańskie, są wyznaczane z pomiarów z odpowiednimi niepewnościami, a także z uwzględnieniem korelacji między błędami poszczególnych zmiennych. Całkowita informacja o niepewnościach pomiaru i współczynnikach korelacji zawarta jest w macierzy  $U_z$ , a uogólniona wektorowa zmienna  $Z$  zawiera wektory



Rys. 2. Charakterystyki funkcji  $\phi(a)$  oraz  $\phi(b)$  dla danych pomiarowych z tab. 1 bez uwzględnienia korelacji współrzędnych w punktach pomiarowych

Tab. 2. Wyznaczone numerycznie parametry prostej regresji bez i uwzględnieniem korelacji

Parametr	Rozwiązanie numeryczne WTLS bez uwzględnienia korelacji	Rozwiązanie numeryczne WTLS z uwzględnieniem korelacji (0,2)
$a / (^{\circ}\text{C} / ^{\circ}\text{C})$	0,9855	0,9844
$b / ^{\circ}\text{C}$	0,7752	0,8550
$u_a / (^{\circ}\text{C} / ^{\circ}\text{C})$	0,0121	0,0112
$u_b / ^{\circ}\text{C}$	0,8189	0,7704
$\rho_{ab}$	-0,9169	-0,9189

Tab. 3. Porównanie metody numerycznej z wynikami analitycznych obliczeń dla rozwiązań regresji prostej  $y$  względem  $x$  i regresji odwrotnej  $x$  względem  $y$

Parametr	Dominacja macierzy $U_y$ rozwiązanie analityczne $\beta^{-1} = \Theta_{n \times n}$	WTLS rozwiązanie numeryczne $U_y \rightarrow \Theta_{n \times n}$ $\beta^{-1} \rightarrow \Theta_{n \times n}$	Dominacja macierzy $U_x$ rozwiązanie analityczne $\beta = \Theta_{n \times n}$	WTLS rozwiązanie numeryczne $U_x \rightarrow \Theta_{n \times n}$ $\beta = \Theta_{n \times n}$
$a / (^{\circ}\text{C} / ^{\circ}\text{C})$	1,0042	1,0042	0,9752	0,9752
$b / ^{\circ}\text{C}$	-0,2594	-0,2594	1,5372	1,5370
$u_a / (^{\circ}\text{C} / ^{\circ}\text{C})$	0,0045	0,0045	0,0089	0,0089
$u_b / ^{\circ}\text{C}$	0,2194	0,2194	0,6760	0,6760
$\rho_{ab}$	-0,9055	-0,9055	-0,9242	-0,9280

Tab. 4. Niepewności rozszerzone wyznaczone w punktach pomiarowych

Lp.	$x_i / ^{\circ}\text{C}$	Bez uwzględnienia korelacji $y_i / ^{\circ}\text{C}$	Uwzględnienie korelacji (0,2) $y_i / ^{\circ}\text{C}$	Dominacja macierzy $U_x$ $y_i / ^{\circ}\text{C}$	Dominacja macierzy $U_y$ $y_i / ^{\circ}\text{C}$
1	30	30,34 ± 1,62	30,39 ± 1,53	29,26 ± 1,40	29,87 ± 0,36
2	50	50,05 ± 1,14	50,07 ± 1,08	48,76 ± 1,00	49,95 ± 0,31
3	70	69,76 ± 1,08	69,76 ± 1,00	68,26 ± 0,82	70,04 ± 0,47
4	90	89,47 ± 1,49	89,45 ± 1,35	87,77 ± 0,99	90,12 ± 0,71
5	110	109,18 ± 2,10	109,14 ± 1,92	107,27 ± 1,39	110,21 ± 0,97

współrzędnych punktów pomiarowych  $\mathbf{X}$  i  $\mathbf{Y}$ . Zaproponowane podejście umożliwia zredukowanie dość złożonego zagadnienia numerycznej minimalizacji funkcji kryterialnej, które realizowane było w różny sposób, do numerycznego wyznaczenia minimum jedno argumentowej funkcji zależnej od współczynnika kierunkowego prostej regresji. W przypadku gdy jedna z macierzy  $U_x$  lub  $U_y$  jest bliska zeru, proponowana metoda numeryczna WTLS, jak wynika z przedstawionych wyników symulacji, jest zbieżna do klasycznych rozwiązań analitycznych dla

regresji prostej  $y$  względem  $x$  i regresji odwrotnej  $x$  względem  $y$ . Na przykładzie wzorcowania pirometru radiacyjnego wyznaczono tą metodą niepewności rozszerzone nieskorelowanych i skorelowanych współrzędnych pomiarowych, przy czym zawężenie korytarza niepewności występuje dla wielkości skorelowanych.

## Bibliografia

- [1] D. York, N. M. Evensen, M. L. Martinez, J. De Basabe Delgado: Unified equations for the slope, intercept and standard errors for the best straight line. *American Journal of Physics*, vol. 72 (2004), s. 367-375.
- [2] N. R. Draper, H. Smith: *Applied Regression Analysis*. 3rd Edition, Wiley, New York 1998.
- [3] J. Puchalski: A new algorithm for generalization of least square method for straight line regression in Cartesian system for fully correlated both coordinates. *International Journal of Automation, Artificial Intelligence and Machine Learning*, vol. 2 (2021), s. 20-54.