

Leszek MOSZCZYŃSKI
Stefan WÓJTOWICZ
Jan SIKORA
Stefan F. FILIPOWICZ

ALGORYTM REGRESJI LINIOWEJ DLA PRZYPADKU NIEPEWNOŚCI OBU ZMIENNYCH W TOMOGRAFII IMPEDANCYJNEJ

STRESZCZENIE *Sukcesem obrazowania obiektu w technikach tomograficznych jest jak największa precyzja poszczególnych pomiarów. Dla zapewnienia dokładności danych pomiarowych tomografii impedancyjnej powinny być uwzględniane błędy poszczególnych pomiarów, zwłaszcza wtedy, gdy różnią się one między sobą istotnie. W impedancyjnej tomografii komputerowej występują przypadki, kiedy obie zależne lub niezależne zmienne pomiarowe obciążone są błędami. W tych przypadkach nie powinna być stosowana standardowa metoda regresji. Do tego typu obliczeń regresji polecany jest zmodyfikowany algorytm Williamsona. W artykule przedstawiono efektywną metodę obliczeń prostej regresji rozszerzając ją na przypadek korelacji błędów obu zmiennych X i Y . Metodę zilustrowano przykładem badania kąta przesunięcia fazowego w zależności od częstotliwości sygnału pobudzającego w układzie wanny pomiarowej.*

Słowa kluczowe: tomografia impedancyjna, regresja liniowa

dr inż. Leszek MOSZCZYŃSKI

e-mail: l.moszczyński@wip.pw.edu.pl

Instytut Technologii Materiałowej, Politechnika Warszawska

dr inż. Stefan WÓJTOWICZ

e-mail: s.wojtowicz@iel.waw.pl

Zakład Metrologii i Badań Nieniszczących, Instytut Elektrotechniki

prof. dr hab. Jan SIKORA

e-mail: sik@iem.pw.edu.pl

dr hab. inż. Stefan F. FILIPOWICZ

e-mail: 2xf@nov.iem.pw.edu.pl

Instytut Elektrotechniki Teoretycznej i Systemów Informatycznych-Pomiarowych,
Politechnika Warszawska

1. WSTĘP

W zagadnieniach tomografii impedancyjnej często zachodzi konieczność powtarzania pomiarów po określonych okresach czasu (w celach porównawczych). Dotyczy to w szczególności monitorowania tkanek lub części ciała podejrzewanych o zmiany chorobowe. W czasie tych, nieraz wielomiesięcznych okresów, parametry samego urządzenia pomiarowego zmieniają się, również wyniki obarczone są określoną niepewnością. Do rozwiązywania tych problemów stosowane są metody regresyjne.

W zagadnieniach impedancyjnej tomografii komputerowej, operujemy danymi obciążonymi niepewnościami pomiaru. W przypadku związku liniowego pomiędzy zmiennymi, ich uwzględnienie wymaga zastosowania jednej z metod regresji ważonej. Są to metody:

1. liniowej regresji ważonej, dla której czynione jest założenie, że odchylenia standardowe wszystkich punktów pomiarowych są stałe a ich poszczególne wagi $w(X_i)=1/\sigma_{X_i}^2$ i $w(Y_i)=1/\sigma_{Y_i}^2$, są takie, że $w(X_i)/w(Y_i) = \lambda$.
2. liniowej regresji ważonej z niezależnymi od siebie błędami dla zmiennych obu współrzędnych X_i i Y_i .

Ten ostatni właśnie przypadek będzie przedmiotem dalszych rozważań.

2. OGÓLNY MODEL REGRESJI

Założmy na wstępie, że posiadane doświadczenie i ogólna znajomość badanego obiektu metodami tomograficznymi, pozwalają zmierzyć m wartości go charakteryzujących. Te zmierzone wartości charakteryzujące obiekt oznaczmy jako $X=(X_1, \dots, X_m)^T$. Prawdziwe ich wartości (bez błędów) mogą być estymowane jako $y = (y_1, \dots, y_m)^T$ ze standardowymi niepewnościami $u(y_i)$ $i=1, \dots, m$.

Jeśli $r(y_i, y_j)$ oznacza współczynnik korelacji między estymowanymi wartościami to kowariancję związaną z estymatorami $U(y_i, y_j)=u(y_i) r(y_i, y_j) u(y_j)$ należy rozpatrywać jako dodatkowy składnik niepewności.

Dodatkowy składnik niepewności opisuje macierz

$$\Sigma = u(y_i y_j^T) = \begin{pmatrix} u^2(y_1) & u(y_1 y_2) & \cdot & \cdot & \cdot & u(y_1 y_m) \\ u(y_2 y_1) & u^2(y_2) & \cdot & \cdot & \cdot & u(y_2 y_m) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u(y_m y_1) & u(y_m y_2) & \cdot & \cdot & \cdot & u^2(y_m) \end{pmatrix} \quad (1)$$

Założmy, że wektory y i X związane są z sobą fizykalnymi lub empirycznymi zależnościami i możliwe jest ustalenie dla $k \geq 0$ parametrów $\beta = (\beta_1, \dots, \beta_k)^T$ takich, że

$$f(\beta, X) = \begin{pmatrix} f_1(\beta_1, X_1) \\ f_1(\beta_2, X_2) \\ \cdot \\ \cdot \\ \cdot \\ f_n(\beta_n, X_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad k \leq n \leq m + k \quad (2)$$

W celu uproszczenia rozważań przyjmiemy, że $f_i : \Omega \longrightarrow R$, $i = 1, \dots, n$ są różniczkowalne. Estymacja zmiennej y_i , macierzy kowariancji Σ i n -wymiarowej funkcji $f(\beta, X)$ jest możliwa różnymi metodami.

W artykule wykorzystano do tego celu metodę najmniejszych kwadratów w której funkcja celu χ^2 (*chi-squared*) zdefiniowana jest jako:

$$\chi^2(X, y) = (y - X)^T \Sigma^{-1} (y - X) \quad (3)$$

przy warunkach $f(\beta, X) = 0$.

Zapisując $\bar{\beta}, \bar{X}$ jako estymaty β, X problem minimalizacji funkcji (4) sprowadza się do poszukiwania rozwiązania równania:

$$\nabla_{(\beta, X, \lambda)} \Phi(\bar{\beta}, \bar{X}, \lambda, y) = 0, \quad (4)$$

gdzie

$$\Phi(\bar{\beta}, \bar{X}, \lambda, y) = (y - X)^T \Sigma^{-1} (y - X) + 2\lambda^T f(\beta, X) \quad (5)$$

dla współczynników Lagrange'a $\lambda = (\lambda_1, \dots, \lambda_n)^T$.

Szczegółowe rozwiązanie tego problemu ilustruje algorytm Yorka [3].

3. METODYKA OBLICZEŃ W REGRESJI WAŻONEJ

Punktem wyjścia do wyprowadzenia algorytmu Yorka jest minimalizacja funkcji kryterialnej (zaproponowanej przez Deminga [2])

$$\chi = \sum_{i=1}^n \{w(X_i)(x_i - X_i)^2 + w(Y_i)(y_i - Y_i)^2\} \quad (6)$$

gdzie: X_i, Y_i są obserwowalnymi wartościami zmiennych $i = 1, \dots, n$, x_i, y_i to estymowane wartości odpowiadające zmiennym X_i i Y_i , $w(X_i), w(Y_i)$ są wagami związanymi z niepewnością X_i i Y_i określenia tak, że:

$$w(X_i) = \frac{1}{\sigma_{X_i}^2}, \quad w(Y_i) = \frac{1}{\sigma_{Y_i}^2}, \quad (7)$$

($\sigma_{X_i}^2, \sigma_{Y_i}^2$ są wariancjami obu obserwowalnych zmiennych).

Zakłada się, że estymowane wartości x_i i y_i związane są zależnością:

$$y_i = ax_i + b \quad \text{dla } i = 1, \dots, n. \quad (8)$$

Jeśli x_i, y_i, a i b mają minimalizować funkcję (6) to spełniony musi być warunek:

$$\delta\chi = \sum_{i=1}^n \{w(X_i)(x_i - X_i)^2 \delta x_i + w(Y_i)(y_i - Y_i)^2 \delta y_i\} = 0, \quad (9)$$

$$a\delta x_i - \delta y_i + \delta b + x_i\delta a = 0. \quad (10)$$

Mnożąc (10) dla wszystkich x_i i y_i przez współczynnik Lagrange'a λ_i i dodając do (9) mamy:

$$\begin{aligned} & \sum_{i=1}^n \delta x_i [w(X_i)(x_i - X_i) + a\lambda_i] + \\ & \sum_{i=1}^n \delta y_i [w(Y_i)(y_i - Y_i) - \lambda_i] + \delta b \sum_{i=1}^n \lambda_i + \delta a \sum_{i=1}^n \lambda_i x_i = 0. \end{aligned} \quad (11)$$

Z równania (11) wynika, że:

$$x_i - X_i = \frac{-a\lambda_i}{w(X_i)}; \quad y_i - Y_i = \frac{\lambda_i}{w(Y_i)}, \quad (12)$$

$$\sum_{i=1}^n \lambda_i = 0 \quad \text{oraz} \quad \sum_{i=1}^n \lambda_i x_i = 0. \quad (13)$$

Wstawiając (12) do (8) otrzymujemy:

$$Y_i + \frac{\lambda_i}{w(Y_i)} = a\left(X_i - \frac{\lambda_i a}{w(X_i)}\right), \quad (14)$$

skąd po przekształceniach:

$$\lambda_i = W_i(aX_i + b - Y_i), \quad (15)$$

$$W_i = \frac{w(X_i)w(Y_i)}{a^2 w(Y_i) + w(X_i)}. \quad (16)$$

Warto podkreślić, że tak zdefiniowana ogólna waga W_i jest wykorzystywana także i w innych algorytmach.

Ponieważ $\sum_{i=1}^n \lambda_i x_i = 0$ zatem $\sum_{i=1}^n W_i(aX_i + b - Y_i)x_i = 0$ i równanie (13) można, zatem zapisać jako:

$$\sum_{i=1}^n W_i(aX_i + b - Y_i)\left(x_i - \frac{a\lambda_i}{w(X_i)}\right) = 0, \quad (17)$$

a wstawiając obliczoną wartość λ_i otrzymamy:

$$\sum_{i=1}^n W_i (aX_i^2 + bX_i - X_i Y_i) - \sum_{i=1}^n \frac{W_i^2 a (aX_i + b - Y_i)^2}{w(X_i)} = 0. \quad (18)$$

Dalsze przekształcenia (18) mogą zostać uproszczone poprzez wprowadzenie średnich ważonych obu zmiennych

$$\bar{Y} = \frac{\sum_{i=1}^n W_i Y_i}{\sum_{i=1}^n W_i}, \quad \bar{X} = \frac{\sum_{i=1}^n W_i X_i}{\sum_{i=1}^n W_i}, \quad (19)$$

stąd

$$\bar{Y} = a\bar{X} + b, \quad (20)$$

$$\sum_{i=1}^n \{W_i (\bar{Y} - a\bar{X}) X_i + aX_i^2 - X_i Y_i\} - \sum_{i=1}^n W_i^2 \frac{a(\bar{Y} - a\bar{X} + aX_i - Y_i)^2}{w(X_i)} = 0. \quad (21)$$

Przyjmując zmienne zastępcze: $U_i = X_i - \bar{X}$ oraz $V_i = Y_i - \bar{Y}$ po przekształceniach otrzymamy:

$$a^3 \sum_{i=1}^n \frac{W_i^2 U_i^2}{w(X_i)} - 2a^2 \sum_{i=1}^n \frac{W_i^2 U_i V_i}{w(X_i)} - a \left\{ \sum_{i=1}^n W_i U_i^2 - \sum_{i=1}^n \frac{W_i^2 U^2}{w(X_i)} \right\} + \sum_{i=1}^n W_i U_i V_i = 0. \quad (22)$$

Równanie (22) nazwane zostało przez Yorka równaniem kubicznym metody najmniejszych kwadratów (ang. *Least-Squares Cubic*).

4. DYSKUSJA ROZWIĄZAŃ RÓWNIANIA KUBICZNEGO

Nieliniowa względem parametru a , złożona postać równania Yorka skłania do rozważenia możliwych (prostszych) przypadków:

a) gdy tylko zmienna niezależna Y_i jest obciążona niepewnością:

W tym przypadku $w(Y_i) = \infty$ i z (16) wynika, że:

$$W_i = \frac{w(\bar{X}_i)}{a^2}. \quad (23)$$

Równanie (22) upraszcza się wówczas do postaci:

$$-a \sum_{i=1}^n W_i U_i^2 + \sum_{i=1}^n W_i U_i V_i w = 0, \quad (24)$$

i ostatecznie mamy:

$$a = \frac{\sum_{i=1}^n w(X_i) U_i V_i}{\sum_{i=1}^n w(Y_i) U_i^2}. \quad (25)$$

Minimalizacja przebiega w kierunku osi $0-X$, a regresja tego typu nazywana jest ważoną regresją x na y (ang. *weighted regression x on y*). Rzeczywiście linia łącząca punkty X_i, Y_i oraz x_i i y_i ma nachylenie:

$$\text{kat } \tau = \frac{y_i - Y_i}{x_i - X_i} = -\frac{1}{a} \frac{w(X_i)}{w(Y_i)}, \quad (26)$$

jak to wynika z równań (12). Jeśli $w(Y_i) = \infty$ to kąt równy jest zero, co potwierdza powyższe spostrzeżenie o położeniu linii „ τ „.

b) tylko zmienna X_i obciążona jest niepewnością:

Oznacza to, że: $w(\bar{W}_i) = \infty$ oraz $W_i = w(Y_i)$ a więc:

$$a = \frac{\sum_{i=1}^n w(X_i) V_i^2}{\sum_{i=1}^n w(Y_i) U_i^2 V_i}. \quad (27)$$

Minimalizacja przebiega w kierunku osi 0-Y, a regresja tego typu nazywana jest ważoną regresją y na x (ang. *weighted regression y on x*).

c) niepewności zmiennych X_i i Y_i spełniają zależność:

$$\frac{w(X_i)}{w(Y_i)} = c, \text{ gdzie } c = \text{const.}$$

Z zależności (2.16) wynika, że dla tego przypadku

$$\text{kąt } l = -c/a. \quad (28)$$

York wykazał, że jeżeli $w(X_i)/w(Y_i) = c$ to:

$$a = \frac{\sum_{i=1}^n w(X_i) V_i^2 - c \sum_{i=1}^n w(X_i) U_i^2}{2 \sum_{i=1}^n w(X_i) U_i V_i} + \frac{\left\{ \sum_{i=1}^n w(X_i) V_i^2 - c \sum_{i=1}^n w(X_i) U_i^2 \right\}^2 + 4c \left(\sum_{i=1}^n w(X_i) U_i V_i \right)^2}{2 \sum_{i=1}^n w(X_i) U_i V_i}^{\frac{1}{2}}. \quad (29)$$

Równanie to jest identyczne z równaniem uzyskanym przez Deminga [2].

d) zmienne X_i i Y_i przyjmują różne niepewności:

W tym najbardziej złożonym przypadku poszukiwany współczynnik kierunkowy prostej a wyznaczany jest z równania (22). Równanie to Reed [5] zapisuje w prostszej postaci:

$$f(a) = a^3 - 3\alpha a^2 + 3\beta a - \gamma = 0, \quad (30)$$

gdzie:

$$\alpha = \frac{2 \sum_{i=1}^n \frac{W_i^2 U_i V_i}{w(X_i)}}{3 \sum_{i=1}^n \frac{W_i^2 U_i^2}{w(X_i)}}; \quad \beta = \frac{\sum_{i=1}^n \frac{W_i^2 V_i}{w(X_i)} - \sum W_i U_i^2}{3 \sum_{i=1}^n \frac{W_i^2 U_i^2}{w(X_i)}}; \quad \gamma = - \frac{\sum_{i=1}^n W_i U_i V_i}{\sum_{i=1}^n \frac{W_i^2 U_i^2}{w(X_i)}}. \quad (31)$$

Równanie (30) jest tylko pozornie równaniem trzeciego stopnia. Można zauważyć, że α, β, γ są funkcjami współczynnika kierunkowego a poprzez wagi W_i . York bez dowodu wskazuje, że równanie (30) ma trzy rzeczywiste pierwiastki. Jednak dowodzi się [5], że równanie to posiada również pierwiastki urojone. Mogą one pojawiać się i ginąć w kolejnych iteracjach.

Problem powyższy rozwiązał Williamson [6] dokonując pseudolinearyzacji równania (30). Wprowadzając oznaczenia:

$$u_i = \frac{1}{w(X_i)}, \quad y_i = \frac{1}{w(Y_i)}, \quad (32)$$

oraz:

$$X'_i = X_i - \bar{X} \quad \bar{X} = \frac{\sum_{i=1}^n W_i X_i}{\sum_{i=1}^n W_i} \quad (33)$$

$$Y'_i = Y_i - \bar{Y}, \quad \bar{Y} = \frac{\sum_{i=1}^n W_i Y_i}{\sum_{i=1}^n W_i}$$

równanie (22) przybiera postać:

$$a^3 \sum_{i=1}^n W_i^2 u_i X_i'^2 - 2a^2 \sum_{i=1}^n W_i^2 u_i X_i' Y_i' - a \left\{ \sum_{i=1}^n W_i^2 X_i'^2 - \sum_{i=1}^n W_i^2 u_i Y_i'^2 \right\} + \sum_{i=1}^n W_i X_i' Y_i' = 0 \quad (34)$$

Williamson, a następnie inni autorzy [8,12] podają ostateczną postać równania normalnego regresji bez wskazania sposobu linearyzacji. W tym celu należy równanie (34) zapisać w postaci

$$a^3 \sum_{i=1}^n W_i^2 u_i X_i'^2 - 2a^2 \sum_{i=1}^n W_i^2 u_i X_i' Y_i' - a \sum_{i=1}^n W_i^2 u_i Y_i'^2 = a \sum_{i=1}^n W_i^2 X_i'^2 - \sum_{i=1}^n W_i X_i' Y_i' \quad (35)$$

a ponieważ

$$(v_i + a^2 u_i) W_i = 1 \quad (36)$$

zatem mnożąc prawą stronę (35) przez (36) i porządkując otrzymujemy

$$a^2 \sum_{i=1}^n W_i^2 u_i X_i' Y_i' - a \sum_{i=1}^n W_i^2 (u_i Y_i'^2 - v_i X_i'^2) - \sum_{i=1}^n W_i^2 X_i' Y_i' v_i = 0 \quad (37)$$

Ostatecznie po dalszych już prostych przekształceniach mamy :

$$a \sum_{i=1}^n W_i z_i X_i' = \sum_{i=1}^n W_i z_i Y_i', \quad (38)$$

gdzie:

$$z_i = W_i (v_i X_i' + a u_i Y_i'). \quad (39)$$

Ponieważ $z_i, W_i, \bar{X}, \bar{Y}$ są funkcjami parametru a , zatem do wyznaczenia współczynników prostej regresji trzeba także zastosować iteracyjną metodę obliczeniową.

5. PROPONOWANA MODYFIKACJA ALGORYTMU WILLIAMSONA

Jeżeli pomiary X_i i Y_i są zależne od siebie to może zdarzyć się, że dodatnim odchyleniom X_i od wartości średniej \bar{X} odpowiadają także dodatnie odchylenia Y_i od \bar{Y} . Tym samym kowariancja

$$\text{cov}(X_i, Y_i) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \quad (40)$$

będzie różna od zera i wówczas algorytm Williamsona wymaga modyfikacji.

Funkcję celu (6) przeformułujemy zatem w następujący sposób:

$$\chi^2 = \sum_{i=1}^n w(X_i)(x_i - X_i)^2 + w(Y_i)(y_i - Y_i)^2 +$$

$$- 2s_i w(X_i)w(Y_i)(x_i - X_i)(y_i - Y_i), \quad (41)$$

gdzie:

$$s_i = r / \sqrt{w(X_i)w(Y_i)} ;$$

r – współczynnik korelacji.

Przyjmując $\frac{\partial \chi^2}{\partial x_i} = 0$ mamy:

$$x_i = \frac{w(X_i) + aw(Y_i) \cdot (Y_i - b) - 2aX_i}{w(X_i) + a^2w(Y_i) \cdot -2asw(X_i)w(Y_i)} \quad (42)$$

Waga W_i i - tego punktu wynosi:

$$W_i = w(X_i) + a^2w(Y_i) - 2asw(X_i)w(Y_i), \quad (43)$$

a funkcja celu (41) upraszcza się do postaci:

$$\chi^2 = \sum_{i=1}^n W_i (b + aX_i - Y_i)^2 \quad (44)$$

Rozwiązując układ równań:

$$\frac{\partial \chi^2}{\partial a} = 0 \quad \text{oraz} \quad \frac{\partial \chi^2}{\partial b} = 0 \quad (45)$$

uzyskujemy równanie trzeciego stopnia, a po jego linearyzacji, równanie (46) w którym z_i obliczane jest ze wzoru:

$$z_i = W_i (v_i X_i' + au_i Y_i' - a s_i X_i' - s_i Y_i'). \quad (46)$$

Biorąc $r = 0$ mamy $s = 0$, a zależność (46) staje się tożsama z (39). Można wykazać, że równanie (46) jest zgodne z ostatnio podanym przez Yorka równaniem zmodyfikowanym od powstania pierwowzoru po blisko 40 latach [3]. Prezentowana tu metoda obliczenia jest znacznie prostsza, a wyprowadzenie pozbawione jest zawikłałości i żmudnych przekształceń algebraicznych jak to ma miejsce w pracach [5, 8, 9].

Wariancje estymowanych wartości a i b wyznacza się poprzez obliczanie pochodnej względem wielkości x_i, y_i . Zmodyfikowany algorytm realizowany jest w sześciu krokach:

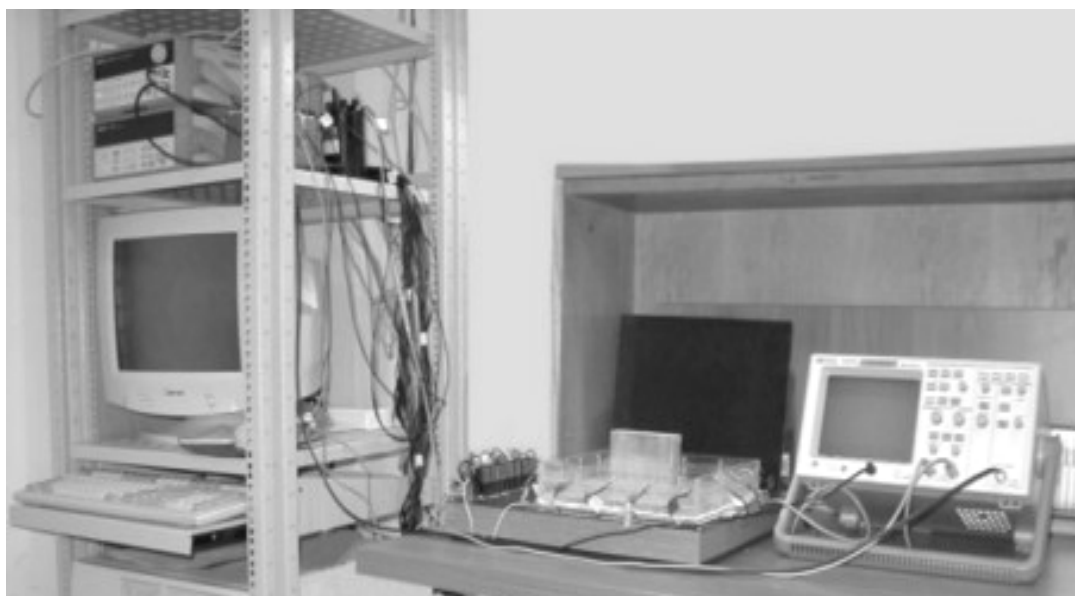
- 1) wstępne obliczenie współczynnika nachylenia prostej regresji a za pomocą metody regresji prostej;
- 2) określenie wag $w(X_i), w(Y_i)$ oraz wagi i -tego punktu W_i z (42) dla ustalonej wartości współczynnika r ;
- 3) na podstawie znajomości X'_i, Y'_i (33) oraz z_i obliczenie parametru a (38) prostej regresji;
- 4) wykorzystując nową wartość a powtórzenie kroków 2 i 3 aż do uzyskania zadowalającej dokładności obliczeń;
- 5) obliczenie parametru b z zależności $b = \bar{Y} - a\bar{X}$;
- 6) dla każdego punktu pomiarowego (X_i, Y_i) ustalenie estymowanej wartości $x_i = \bar{X} + z_i$ oraz $y_i = \bar{Y} + z_i$.

Obliczenia są proste, a algorytm może być zrealizowany także w arkuszu kalkulacyjnym Excel. Efektywność algorytmu można jeszcze zwiększyć dobierając w pierwszym kroku iteracji wartości parametru a z obliczeń metodą regresji prostej.

Dokładność dopasowania prostej regresji może być oceniona poprzez obliczenie wartości χ^2 (11) dla $n-2$ stopni swobody. Działanie algorytmu sprawdzono na danych Pearsona [1] uznawanych za dane testujące dla procedur obliczeniowych regresji.

6. ZASTOSOWANIE ALGORYTMU REGRESJI WAŻONEJ DO WYNIKÓW EKSPERYMENTALNYCH

Na rysunku 1 przedstawiono stanowisko eksperymentalne do badania przesunięcia fazowego w tomografii impedancyjnej.



Rys. 1. Stanowisko pomiarowe do badania przesunięcia fazowego w tomografii impedancyjnej

Pomiary częstotliwości i przesunięcia fazowego obciążone są jak w każdym przyrządzie cyfrowym: błędami dyskretyzacji, błędami wzorca i błędami bramkowania. Błąd dyskretyzacji opisywany rozkładem Simpsona można oszacować operując stopniem wypełnienia licznika przyrządu. Przy małym wypełnieniu licznika błędy dyskretyzacji mogą być znacznie większe niż przy jego zapelnieniu. Błąd bramkowania wynika natomiast z niedokładności określenia momentu początku i końca mierzonego okresu. Zależy on więc także od częstotliwości sygnału. Czynniki te wyjaśniają zróżnicowanie niepewności pomiarów w dla różnych badanych wartości częstotliwości i przesunięć fazowych.

W prezentowanych badaniach wartości wariancji częstotliwości (σ_x) i przesunięć fazowych (σ_{y1} i σ_{y2}) dla każdej mierzonej wartości określono eksperymentalnie. Uwzględniono horyzont długoczasowy tj. zmienność wskazań przyrządów po czasie dłuższym niż miesiąc. Pomiar kąta dokonywany był oscyloskopem HP 34600A a wybrane uśrednione wyniki pomiaru przedstawiono w tab. 1.

Wykorzystując opisywany algorytm przeprowadzono obliczenia dla różnych niepewności pomiarów. W tabeli 2 i 3 przedstawiono dziesięć przypadków doboru współczynników regresji dotyczących charakterystyk przesunięć fazowych sygnału tomografii impedancyjnej przy rozpoznawaniu obiektów w postaci klocka metalowego i połowy jabłka.

TABELA 1

Wyniki pomiaru przesunięcia fazowego przy różnej częstotliwości pobudzenia

Pomiar częstotliwości generatora		Pomiar przesunięcia fazowego obiektu			
Częstotliwość generatora X	Błąd wskazań σ_x	Klocek aluminiowy		Połowa jabłka	
		Przesunięcie fazowe Y_1	Błąd wskazań σ_{y_1}	Przesunięcie fazowe Y_2	Błąd wskazań σ_{y_2}
kHz	kHz	stopnie	stopnie	stopnie	stopnie
10					
100	0.50	0.35	0.75	-0.55	0.15
200	1.00	1.80	1.10	3.05	0.15
300	1.50	4.45	1.25	7.00	0.50
400	2.00	8.90	1.70	11.55	0.35
500	2.50	11.85	1.65	15.75	0.55
600	1.00	14.95	1.15	19.50	0.60
700	4.00	18.10	1.10	23.00	0.30
800	3.00	21.15	1.15	26.50	0.50
900	7.50	24.00	1.00	29.30	0.20
1000	6.50	26.50	1.50	32.60	0.50
1100	5.00	31.00	1.00	35.65	0.55
1200	7.50	34.00	2.00	38.50	0.50
1300	5.00	33.00	1.00	40.95	0.75
1400	7.50	35.60	1.50	43.09	0.30
1500	7.50	37.20	1.90	45.40	0.50
1600	5.00	38.65	1.45	47.95	0.25
1700	7.50	41.55	1.55	49.95	0.25
1800	7.50	45.50	1.50	51.10	0.60
1900	5.00	50.00	0.10	52.40	0.30
2000	10.00	51.15	0.75	54.50	0.50

TABELA 2

Wyniki obliczeń współczynników regresji dla różnych wariacji danych pomiarowych i metalowego prostopadłościanu

Nr	Przyjmowane niepewności		Element aluminiowy prostopadłościenny			
			Współczynniki regresji		Niepewności współczynników	
	a	b	s_a	s_b		
1	Algorytm Williamsona	Algorytm William.	0.02616	-0.26169	1.3533e-004	0.00135
2	$u_i = \text{eps}$	v_i	0.02617	-0.2617464	1.338581e-004	0.00133
3	u_i	$v_i = \text{eps}$	0.02568	-0.256898	1.160153e-005	1.16015 e-004
4	$u_i = 0$	$v_i = 0$	0.02643	-1.2258486	5.826052e-004	0.36149
5	$u_i = \text{max}$	$v_i = \text{eps}$	0.02670	-1.4954083	3.062929e-005	0.03575
6	$u_i = \text{min}$	$v_i = \text{eps}$	0.02670	-1.4954083	6.848919e-006	0.00799
7	$u_i = \text{min}$	v_i	0.02617	-0.2642479	1.345014e-004	0.01866
8	$u_i = \text{max}$	v_i	0.02619	-0.3098410	1.459956e-004	0.08230
9	u_i	$v_i = \text{max}$	0.02643	-1.2273105	5.107459e-004	0.59686
10	u_i	$v_i = \text{min}$	0.02644	-1.2300423	3.13213e-004	0.36576

esp – najmniejsza ze zmiennoprzecinkowych liczb akceptowalna przez algorytm ($\text{esp} = 2.22 \cdot 10^{-16}$) –
 – dotyczy tabel 2 i 3.

TABELA 3

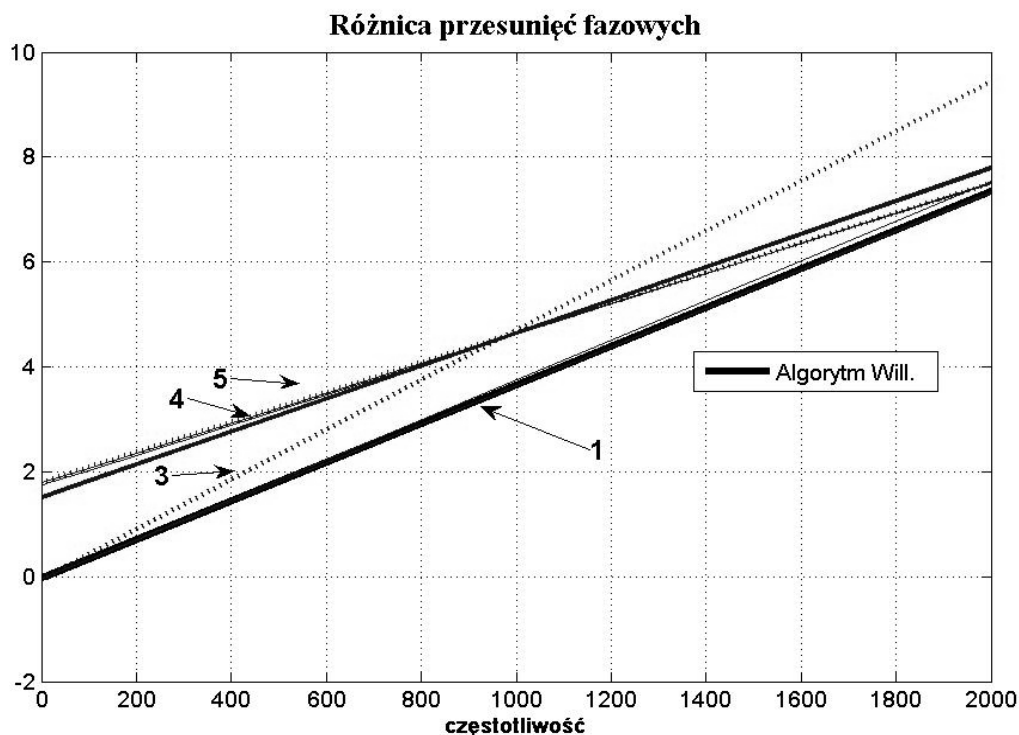
Wyniki obliczeń współczynników regresji dla różnych wariacji danych pomiarowych i połowy jabłka

Nr	Przyjmowane niepewności		Połowa jabłka			
			Współczynniki regresji		Niepewności współczynników	
			a	b	s_a	s_b
1	Algorytm Williamsona	Algorytm Williamsona	0.02985	-0.2985	1.16116e-004	0.00116
2	$u_i = \text{eps}$	v_i	0.02985	-0.2985	1.15113e-004	0.00115
3	u_i	$v_i = \text{eps}$	0.03042	-0.3042	1.37552e-005	1.37552e-004
4	$u_i = 0$	$v_i = 0$	0.02928	0.57053	8.82698e-004	0.54769
5	$u_i = \text{max}$	$v_i = \text{eps}$	0.02984	0.01197	3.43737e-005	0.04008
6	$u_i = \text{min}$	$v_i = \text{eps}$	0.02984	-0.0119	7.68620e-006	0.00896
7	$u_i = \text{min}$	v_i	0.02986	-0.3026	1.16273e-004	0.02122
8	$u_i = \text{max}$	v_i	0.02999	-0.3628	1.34083e-004	0.08895
9	u_i	$v_i = \text{max}$	0.02929	0.55768	3.13569e-004	0.36608
10	u_i	$v_i = \text{min}$	0.02933	0.50619	1.42105e-004	0.16503

Wiersz pierwszy tabeli 1 i 2 przedstawia wyniki uzyskane przy zastosowaniu opisanego zmodyfikowanego algorytmu Williamsona. Pozostałe przypadki dotyczą obliczeń z uwzględnieniem różnych możliwych kombinacji wariacji zmiennych X i Y . Tak więc na przykład zapis u_i w kolumnie drugiej, a w kolumnie trzeciej $v_i = \text{eps}$ oznacza, że pomiar częstotliwości obciążony był w każdym punkcie pomiarowym wariacją u_i (z tab. 1), natomiast błędy przesunięć fazowych przyjęto jednakowe dla wszystkich pomiarów o wartości $\text{eps} = 2,22 \cdot 10^{-16}$. Przypadek $u_i = 0$ oraz $v_i = 0$ oznacza, że do obliczeń użyto standardowej metody regresji OLS (ang. *Ordinary Least Squares*) dostępny w większości pakietów obliczeniowych.

Analizując wielkości w tabelach 2 i 3 zauważamy, że stosując algorytm OLS uzyskujemy dobór parametrów regresji obciążony bardzo dużą niepewnością parametru b , np. dla rozpoznawania połówki jabłka odchylenie standardowe $s_b = 0.54769$. Przyjmując niepewności poszczególnych pomiarów nawet w bardzo wąskim zakresie (dla częstotliwości $u_i = \text{min}$, a dla kąta $v_i = \text{eps}$) otrzymujemy znacznie lepsze dopasowanie prostej regresji. Współczynniki a regresji obliczane różnymi metodami są bliskie sobie. Ich waga zaczyna jednak być znacząca przy dużych częstotliwościach. Przy niskich częstotliwościach natomiast decydujące znaczenie odgrywa współczynnik b . Tu różnice dla poszczególnych przypadków są duże.

Na rysunku 2 pokazano przebieg zależności różnicy przesunięć fazowych w funkcji częstotliwości dla przedmiotu metalowego i połowy jabłka uzyskane dla wyników podanych w tab. 2 i 3.



Rys. 2. Różnice przesunięć fazowych przy różnych metodach obliczeń
Oznaczenia prostych: opis w tekście

Prosta oznaczona cyfrą 1 obliczona została według prezentowanego wcześniej algorytmu. Widoczne jest, że dla częstotliwości równej zero i uzyskane przesunięcie fazowe jest zerowe. Świadczy to o poprawności doboru prostej, czego nie można powiedzieć o pozostałych prostych (obliczonych według wierszy 3, 4, 5 z tab. 1 i 2). Również nachylenie prostej informuje, że mamy do czynienia z obiektem o charakterze rezystancyjno-pojemnościowym; umożliwia to wyznaczenie wzajemnych relacji między tymi elementami, co jest szczególnie ważne w budowaniu modelu matematycznego obiektu oraz może być wykorzystane w badaniach symulacyjnych (obwodowych) np. przy użyciu języka NAP.

Znaczenie właściwego doboru współczynników regresji najlepiej można ocenić porównując ze sobą charakterystyki przesunięć fazowych różnych lub tych samych obiektów, lecz wykonanych w różnych okresach czasu.

7. PODSUMOWANIE I WNIOSKI

Sukcesem obrazowania obiektu w technikach tomograficznych jest jak największa precyzja poszczególnych pomiarów. Dla zapewnienia dokładności

danych pomiarowych tomografii impedancyjnej powinny być uwzględniane błędy poszczególnych pomiarów, zwłaszcza wtedy, gdy różnią się one między sobą istotnie. Do tego typu obliczeń regresji polecany jest zmodyfikowany algorytm Williamsona. Jest on prostszy niż stosowany algorytm Yorka i może być zaimplementowany także w programie Excel. Przykład zastosowania tej metody do badań tomograficznych pokazuje, że nieuwzględnienie błędów w obu wielkościach mierzonych prowadzi do znacznego poszerzenia przedziału ufności prostej kalibracji. Może to wpływać znacząco na ocenę wyników badań tomograficznych.

Zmodyfikowany algorytm Williamsona należy stosować zwłaszcza dla przypadków, kiedy w celach porównawczych prowadzone badania są powtarzane po długim okresie czasowym.

LITERATURA

1. Pearson K., On lines and planes of closest fit to systems of points in space, *Phil. Mag.*, 2 str. 912-916, 1901.
2. Deming W.E., *Statistical Adjustment of Data*, Wiley, New York, 1964.
3. York D., Least-squares fitting of straight line, *Can. J. Phys.*, 44, str. 1079-1086, 1966.
4. Jeffrys W.H., Robust estimation when more than one variable per equation of condition has error, *Biometrika*, 77, str. 597-607, 1990.
5. Reed B.C., Linear least squares fits with errors in both coordinates, *Am. J. Phys.*, 57, n.189, str. 642-646, 1990.
6. Williamson J.H., Least-squares fitting of a straight line, *Can. J. Phys.*, 46, str. 1845-1847, 1968.
7. York D., Unified equation for slope, intercept, and standard errors of the best straight line, *Am. J. Phys.*, 72 n.3, str. 367-275, 2004.
8. Cecchi G.C., Error analysis of the parameters of a least-squares determined curve when both variables have uncertainties, *Meas. Sci. Technol.*, 2, str. 1127-1129, 1990.
9. Lybanon M., A better least-squares method when both variables have uncertainties, *Am. J. Phys.*, 52 n.1, str. 22-26, 1984.
10. Orear J., Least squares when both variables have uncertainties, *Am. J. Phys.*, 50 n.10, str. 912-915, 1982.
11. Erratum for "Least squares when both variables have uncertainties, *Am. J. Phys.*, 52 n.3, str. 278-279, 1984.
12. Riu J., Rius F.X., Method comparison using regression with uncertainties in both axes, *Trends in analytical chemistry*, n. 16, str. 211-216, 1997

Rękopis dostarczono, dnia 20.02.2007 r.

Opiniował: prof. dr hab. inż. Antoni Cieśla

LINEAR REGRESSION ALGORITHM
IN CASE OF UNCERTAINTY BOTH VARIABLES
IN ELECTRICAL IMPEDANCE TOMOGRAPHY

Leszek MOSZCZYŃSKI, Stefan WÓJTOWICZ,
Jan SIKORA, Stefan F. FILIPOWICZ

ABSTRACT *For Electrical Impedance Tomography often we have to deal with the case when both variables – independent and dependent ones are collected with a certain error level. In order to achieve better results the standard regression method should not be applied in such cases. New Linear Regression Method expanded on the case of error correlation for both variables X and Y is presented in this paper. Electrical Impedance Tomography (EIT) problems very often need to be repeated after some period of time. It is necessary for tissue monitoring or even for some parts of the human body monitoring, which are changed by a certain illness. During this very often, long periods of time parameters of measuring system also are changed, that is why the output variables have a certain level of uncertainty. In order to solve such a problem, usually regression method is employed. This method is illustrated by an example of phase shift angle investigation with respect to the excitation signal of measuring bath.*