

Adam G. POLAK, Janusz MROCZKA

POLITECHNIKA WROCŁAWSKA, KATEDRA METROLOGII ELEKTRONICZNEJ I FOTONICZNEJ

Regularyzacja identyfikacji obiektów złożonych opisanych modelami nieliniowymi

Dr inż. Adam G. POLAK

Adiunkt w Katedrze Metrologii Elektronicznej i Fotonicznej na Wydziale Elektroniki Politechniki Wrocławskiej. Zajmuje się modelowaniem i identyfikacją obiektów złożonych oraz nieinwazyjnymi metodami pomiaru mechanicznych właściwości układu oddechowego.



e-mail: Adam.Polak@pwr.wroc.pl

Prof. dr hab. inż. Janusz MROCZKA

Kierownik Katedry Metrologii Elektronicznej i Fotonicznej Politechniki Wrocławskiej. Zajmuje się metodologią obserwacji i eksperymentu, algorytmizacją problemu odwrotnego, modelowaniem matematycznym pól fizycznych, analizą spektralną i polaryzacyjną promieniowania rozproszonego, reprezentacjami czasowo-częstotliwościowymi w przetwarzaniu danych.



e-mail: Janusz.Mroczka@pwr.wroc.pl

Streszczenie

W referacie przedstawiono modyfikację regularyzacji Tichonowa-Phillipsa (TP) przystosowującą ją do estymacji parametrów modeli nieliniowych źle uwarunkowanych numerycznie. Zaproponowane podejście porównano z metodami Gaussa-Newtona (GN), Levenberga-Marquardta (LM) oraz metodą łączącą LM z TP (LMTP). Analizę właściwości czterech zaimplementowanych algorytmów przeprowadzono metodą Monte Carlo. Pokazała ona, że w przypadku identyfikacji modeli nieliniowych zawierających parametry słabo określone przez dane pomiarowe i jednocześnie charakteryzujące się „regularnym” rozkładem wartości w wektorze parametrów, najlepsze wyniki daje estymacja metodą Tichonowa-Phillipsa.

Słowa kluczowe: regularyzacja estymacji, algorytm Levenberga-Marquardta, metoda Tichonowa-Phillipsa

Regularized identification of complex objects described by nonlinear models

Abstract

In the paper a modification of the Tikhonov-Phillips regularisation enabling the estimation of parameters of numerically ill-conditioned nonlinear models is presented. This approach was compared with the Gauss-Newton (GN) and Levenberg-Marquardt (LM) methods, as well as with a method combining LM with TP one (LMTP). The analysis of the four computer-implemented algorithms was done by the Monte Carlo simulations. They have shown that the result of identification of a nonlinear model possessing weakly defined, however “regularly” distributed parameters, is the most accurate when using the Tikhonov-Phillips method.

Keywords: regularized estimation, Levenberg-Marquardt algorithm, Tikhonov-Phillips method

1. Wprowadzenie

Wiele z otaczających człowieka obiektów ma wysoce złożoną strukturę i cechuje się wielością, różnorodnością i nieliniowością zachodzących w nich zjawisk. Podstawowym narzędziem poznania ich budowy i działania są pomiary, w szczególności pomiary pośrednie wykorzystujące modele matematyczne, które wiążą rejestrowane sygnały z wewnętrznymi właściwościami obiektów. Tak realizowane pomiary pośrednie sprowadzają się do identyfikacji parametrycznej wykorzystującej algorytmy numeryczne. Złożoność struktury badanych obiektów i nieliniowości zachodzących w nich zjawisk powodują, że estymacja parametrów jest zazwyczaj źle uwarunkowana numerycznie.

W celu poprawy uwarunkowania numerycznego stosowane są metody regularyzacji, wykorzystywane przede wszystkim w zadaniach odtwarzania sygnałów pomiarowych i analizie danych wielowymiarowych. Sama nazwa operacji nawiązuje do nadania poszukiwanemu rozwiązaniu bardziej „regularnej” formy w przeciwieństwie do efektu uzyskiwanego bez niej, kiedy to rozwiązanie jest zdominowane przez propagujący i wzmocniony szum pomiarowy. Regularyzacja estymacji parametrów modeli obiektów fizycznych jest rzadziej spotykana, a do najpopularniejszych metod należą regresja grzbietowa (RR) i regresja składników głównych (PCR). Metody te jednak zakładają liniowość modelu pod względem parametrów.

Na początku niniejszego referatu przedstawiono metodę Levenberga-Marquardta służącą regularyzacji identyfikacji nieliniowych modeli obiektów złożonych. Następnie zaproponowano zastosowanie metody Tichonowa-Phillipsa w przypadku identyfikacji modeli o parametrach skupionych opisujących przestrzenny (lub inny) rozkład uwzględnionej wielkości fizycznej. Ponieważ oryginalna metoda Tichonowa-Phillipsa jest liniowa, dokonano jej modyfikacji umożliwiającej estymację nieliniową. Dodatkowo wyprowadzony został wzór na zregularyzowaną estymację parametrów modeli nieliniowych łączącą w sobie metody Levenberga-Marquardta i Tichonowa-Phillipsa. Przedstawiono też badania symulacyjne demonstrujące efekt zastosowania omówionych metod do estymacji parametrów dobrego specjalnie w tym celu modelu nieliniowego.

2. Metody regularyzacji estymacji

Przypadki złego uwarunkowania numerycznego estymacji mają miejsce w sytuacjach, gdy wyznacznik odwracanej macierzy informacyjnej jest bliski zeru, co przekłada się na bardzo dużą wartość liczby uwarunkowania, czyli stosunku jej największej wartości osobliwej do wartości najmniejszej. Obliczane wtedy estymaty są wyjątkowo wrażliwe na zakłócenia zawarte w danych pomiarowych. Regularyzacja polega na zastąpieniu zagadnienia oryginalnego problemem przybliżonym, którego rozwiązanie jest znacząco mniej wrażliwe na błędy.

Można wyróżnić kilka podstawowych mechanizmów uwzględnianych w algorytmach numerycznych i dających efekt regularyzacji odtwarzania sygnałów [8]. Należą do nich: dobór przedziału dyskretyzacji ciągłego modelu matematycznego, metody wariacyjne, parametryzacja odtwarzanych sygnałów, uwzględnienie wielkości charakteryzujących poziomy niepewności, wcześniejsze zakończenie działania algorytmów iteracyjnych czy zastosowanie ograniczeń w dziedzinie transformacji rekonstruowanego sygnału.

Typowe podejście do regularyzacji estymacji polega na uzupełnieniu minimalizowanej funkcji kryterialnej V o czynnik kary Ω zależny od wektora parametrów. Daje to następującą postać zregularyzowanej funkcji kryterialnej:

$$V_R(\boldsymbol{\theta}) = V(\boldsymbol{\theta}) + k\Omega(\boldsymbol{\theta}), \quad (1)$$

gdzie k jest skalarnym współczynnikiem regularyzacji. Zaletą regularyzacji jest poprawa identyfikowalności modelu bez potrzeby modyfikacji jego struktury, a wadą wzrost złożoności obliczeniowej i pojawienie się obciążenia estymatorów. Minimalizując $V_R(\boldsymbol{\theta})$ uzyskuje się obciążony estymator

$$\hat{\boldsymbol{\theta}}_R = \arg \min_{\boldsymbol{\theta}} V_R(\boldsymbol{\theta}), \quad (2)$$

którego obciążenie \mathbf{b}_R można oszacować linearyzując V_R poprzez rozwinięcie w szereg Taylora. Dane jest ono zależnością [4]:

$$\mathbf{b}_R = -k(\mathbf{H}_V + k\mathbf{H}_\Omega)^{-1} \nabla_{\boldsymbol{\theta}} \Omega, \quad (3)$$

gdzie macierze hessianu \mathbf{H} i gradient ∇ są określone następująco:

$$\mathbf{H}_V = \frac{\partial^2 V}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}}, \quad \mathbf{H}_\Omega = \frac{\partial^2 \Omega}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}}, \quad \nabla_{\boldsymbol{\theta}} \Omega = \frac{\partial \Omega}{\partial \boldsymbol{\theta}}. \quad (4)$$

Obciążenie rosnące wraz ze współczynnikiem k jest ceną za uzyskiwane wtedy zmniejszanie się wariancji \mathbf{S}_R . Macierz kowariancji dana jest ogólnym wzorem [4]:

$$\mathbf{S}_R = (\mathbf{H}_V + k\mathbf{H}_\Omega)^{-1} \mathbf{P} (\mathbf{H}_V + k\mathbf{H}_\Omega)^{-1}, \quad (5)$$

$$\mathbf{P} = \mathbf{E}(\nabla_{\boldsymbol{\theta}} V \cdot \nabla_{\boldsymbol{\theta}}^T V).$$

2.1. Metoda Levenberga-Marquardta (LM)

Algorytm zregularyzowanej estymacji zaproponowany przez Levenberga i Marquardta [5], [6], [7] jest w istocie nieliniową wersją regresji grzbietowej – metody stosowanej w przypadku modeli liniowych [2], [3]. Użyty czynnik kary uwzględnia różnicę w L^2 między estymowanym wektorem parametrów $\boldsymbol{\theta}$ a wybranymi wartościami $\boldsymbol{\theta}^*$, którymi mogą być np. parametry typowego punktu pracy systemu (w regresji grzbietowej przyjmuje się $\boldsymbol{\theta}^* = \mathbf{0}$). Wtedy funkcja kryterialna V_{LM} przyjmuje postać:

$$V_{LM}(\boldsymbol{\theta}) = \frac{1}{2} \|\mathbf{y} - \mathbf{y}_m(\boldsymbol{\theta})\|^2 + \frac{1}{2} k \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|^2, \quad (6)$$

gdzie \mathbf{y} jest wektorem danych pomiarowych, a \mathbf{y}_m jest wyjściem modelu. Przyrównanie do zera gradientu zregularyzowanej funkcji kryterialnej połączone z lokalną linearyzacją modelu pozwala na wyprowadzenie wzoru rekurencyjnego, w którym w $i+1$ kroku obliczeń wykorzystuje się dane uzyskane w kroku i :

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i + \mu [\mathbf{X}^T(\hat{\boldsymbol{\theta}}_i) \mathbf{X}(\hat{\boldsymbol{\theta}}_i) + k\mathbf{I}]^{-1} \mathbf{X}^T(\hat{\boldsymbol{\theta}}_i) \boldsymbol{\varepsilon}(\hat{\boldsymbol{\theta}}_i), \quad (7)$$

gdzie parametr μ kontroluje długość kroku, $\mathbf{X} = \partial \mathbf{y}_m / \partial \boldsymbol{\theta}$, a $\boldsymbol{\varepsilon}$ to rezydium. Wartość współczynnika regularyzacji k można zmieniać w każdej iteracji, w zależności od uzyskanego efektu [11].

W metodzie tej, zgodnie ze wzorami ogólnymi, przybliżone zależności na obciążenie wektora estymatorów i jego macierz kowariancji są następujące:

$$\mathbf{b}_{LM} \approx -k(\mathbf{X}^T \mathbf{X} + k\mathbf{I})^{-1} (\boldsymbol{\theta}_0 - \boldsymbol{\theta}^*), \quad (8)$$

$$\mathbf{S}_{LM} \approx \sigma_e^2 (\mathbf{X}^T \mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X} + k\mathbf{I})^{-1}. \quad (9)$$

2.2. Metoda Tichonowa-Phillipsa (TP)

Metoda zaproponowana przez Phillipsa i Twomeya [10], [13], [12] do regularyzacji odtwarzania sygnałów, zwana jest też regularyzacją Phillipsa-Twomeya lub Tichonowa-Millera [11], [9]. W przypadku parametrycznej identyfikacji modelu jej zadaniem będzie zapewnienie „regularności” rozkładu sąsiadujących estymatorów. Czynnik kary otrzymuje się wymnażając operator macierzowy \mathbf{D} przez wektor estymowanych parametrów. Funkcja kryterialna V_{TP} przyjmuje wtedy postać:

$$V_{TP}(\boldsymbol{\theta}) = \frac{1}{2} \|\mathbf{y} - \mathbf{y}_m(\boldsymbol{\theta})\|^2 + \frac{1}{2} \gamma \|\mathbf{D}\boldsymbol{\theta}\|^2. \quad (10)$$

Zazwyczaj macierz \mathbf{D} jest dyskretnym operatorem różniczkowania n -tego stopnia, niezależnym od $\boldsymbol{\theta}$. Wymusza ona „gładkość” estymowanego wektora poprzez minimalizację różnic pomiędzy kolejnymi jego elementami. W przypadku pierwszej pochodnej

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \quad (11)$$

W celu wyprowadzenia wzoru na obliczenia iteracyjne w przypadku zregularyzowanej estymacji parametrów modelu nieliniowego należy dokonać, analogicznie jak w przypadku regresji grzbietowej, linearyzacji modelu. Przyrównując gradient funkcji kryterialnej do zera uzyskuje się wzór na iteracyjne oszacowywanie estymatora parametrów:

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i + \mu [\mathbf{X}^T(\hat{\boldsymbol{\theta}}_i) \mathbf{X}(\hat{\boldsymbol{\theta}}_i) + \gamma \mathbf{D}^T \mathbf{D}]^{-1} \cdot [\mathbf{X}^T(\hat{\boldsymbol{\theta}}_i) \boldsymbol{\varepsilon}(\hat{\boldsymbol{\theta}}_i) - \gamma \mathbf{D}^T \mathbf{D} \hat{\boldsymbol{\theta}}_i]. \quad (12)$$

W metodzie Tichonowa-Phillipsa przybliżone zależności na obciążenie wektora estymatorów i jego macierz kowariancji dane są następująco:

$$\mathbf{b}_{TP} \approx -\gamma (\mathbf{X}^T \mathbf{X} + \gamma \mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{D} \boldsymbol{\theta}_0, \quad (13)$$

$$\mathbf{S}_{TP} \approx \sigma_e^2 (\mathbf{X}^T \mathbf{X} + \gamma \mathbf{D}^T \mathbf{D})^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X} + \gamma \mathbf{D}^T \mathbf{D})^{-1} \quad (14)$$

2.3. Połączenie metod LM i TP

Modyfikację kierunku poszukiwania minimum funkcji kryterialnej, poprawę uwarunkowania numerycznego i „wygładzanie” wektora estymowanych parametrów można uzyskać łącząc opisane wcześniej metody LM i TP. Funkcja kryterialna przyjmuje wtedy postać:

$$V_{LMTP} = \frac{1}{2} \|\mathbf{y} - \mathbf{y}_m(\boldsymbol{\theta})\|^2 + \frac{1}{2} \gamma \|\mathbf{D}\boldsymbol{\theta}\|^2 + \frac{1}{2} k \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|^2. \quad (15)$$

Linearyzując ponownie model i przyrównując gradient funkcji kryterialnej do zera otrzymuje się:

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + \mu \left[\mathbf{X}^T(\boldsymbol{\theta}_i) \mathbf{X}(\boldsymbol{\theta}_i) + \gamma \mathbf{D}^T \mathbf{D} + k \mathbf{I} \right]^{-1} \cdot \left[\mathbf{X}^T(\boldsymbol{\theta}_i) \boldsymbol{\varepsilon}(\boldsymbol{\theta}_i) - \gamma \mathbf{D}^T \mathbf{D} \boldsymbol{\theta}_i \right]. \quad (16)$$

Postać wyrażenia na obciążenie i macierz kowariancji tak zregularyzowanego estymatora jest analogiczna do poprzednich:

$$\mathbf{b}_{LMTP} \approx - \left(\mathbf{X}^T \mathbf{X} + k \mathbf{I} + \gamma \mathbf{D}^T \mathbf{D} \right)^{-1} \cdot \left[k(\boldsymbol{\theta}_0 - \boldsymbol{\theta}^*) + \gamma \mathbf{D}^T \mathbf{D} \boldsymbol{\theta}_0 \right], \quad (17)$$

$$\mathbf{S}_{LMTP} \approx \sigma_e^2 \left(\mathbf{X}^T \mathbf{X} + k \mathbf{I} + \gamma \mathbf{D}^T \mathbf{D} \right)^{-1} \mathbf{X}^T \mathbf{X} \cdot \left(\mathbf{X}^T \mathbf{X} + k \mathbf{I} + \gamma \mathbf{D}^T \mathbf{D} \right)^{-1}. \quad (18)$$

3. Badania symulacyjne algorytmów

Właściwości standardowej metody Gaussa-Newtona (GN), metody Levenberga-Marquardta (LM) i dwóch zaproponowanych nieliniowych wersji algorytmów regularyzacji, tj. metody Tichonowa-Phillipsa (TP) oraz połączonej metody Levenberga-Marquardta i Tichonowa-Phillipsa (LMTP) porównano na drodze symulacji komputerowych przeprowadzonych w środowisku Matlab.

W pierwszej kolejności, wykorzystując odpowiednio dobrany nieliniowy model matematyczny „identyfikowanego obiektu fizycznego”, wygenerowano syntetyczne dane pomiarowe. Struktura modelu, wartości jego parametrów $\boldsymbol{\theta}$ oraz charakter i zakres sygnału wejściowego u powodują, że estymacja parametrów modelu na podstawie zaszumionych danych pomiarowych jest źle uwarunkowana numerycznie. Wyjście modelu y_m opisane jest następująco:

$$y_m = \sum_{j=1}^{10} \alpha^j (\boldsymbol{\theta}_j + u)^j, \quad (19)$$

gdzie współczynnik α nie jest identyfikowanym parametrem lecz znaną stałą. W przedstawionym modelu wartość α ustalono na 0.5, dzięki czemu wpływ parametrów o rosnących numerach na wyjście jest coraz mniejszy i tym samym stają się one coraz trudniej identyfikowalne (pierwszy czynnik sumy przemnożony jest przez 0.5, a dziesiąty przez ok. 0.001). Wejście u narasta liniowo od 0 do 0.5, a dane pomiarowe uzyskano dla przyrostów 0.01, co daje 51 symulowanych punktów pomiarowych. Zakres u i wysokie potęgi przy ostatnich sumowanych składnikach powodują, że wpływy stojących tam parametrów stają się coraz bardziej współliniowe. Ostateczne dane wyjściowe y obliczono sumując wyjście modelu y_m z wygenerowaną realizacją szumu białego o odchyleniu standardowym równym 1% wartości maksymalnej y_m . Ponieważ metoda Tichonowa-Phillipsa predestynowana jest do „wygładzania” różnic estymat parametrów umieszczonych obok siebie w wektorze $\boldsymbol{\theta}$, wartości prawdziwe $\boldsymbol{\theta}_0$ podczas generacji dobrano zgodnie z tą przesłanką.

Po zaimplementowaniu wymienionych algorytmów wyznaczono liczby uwarunkowania numerycznego odwracanych w poszczególnych metodach macierzy, czyli odpowiednio GN: $(\mathbf{X}^T \mathbf{X})$, LM: $(\mathbf{X}^T \mathbf{X} + k \mathbf{I})$, TP: $(\mathbf{X}^T \mathbf{X} + \gamma \mathbf{D}^T \mathbf{D})$ i LMTP: $(\mathbf{X}^T \mathbf{X} + k \mathbf{I} + \gamma \mathbf{D}^T \mathbf{D})$.

Do symulacyjnej oceny algorytmów zastosowano metodę Monte Carlo. W tym celu 10000 razy powtórzono procedurę estymacji parametrów powyższymi metodami (poza GN), stosując wektor startowy $\boldsymbol{\theta}_{sr} = 0.2 \cdot \mathbf{1}$. Za każdym razem generowana

była nowa realizacja szumu białego i sumowana z wyjściem modelu. Dane te używano do obliczenia estymatorów $\boldsymbol{\theta}_{LM}$, $\boldsymbol{\theta}_{TP}$ i $\boldsymbol{\theta}_{LMTP}$. Dla każdej z metod, korzystając z 10000 wektorów estymat i znanych wzorów, oszacowano wartości oczekiwane i odchylenia standardowe estymatorów. Różnice między obliczonymi wartościami średnimi i prawdziwymi użytymi do generacji danych dają oszacowanie obciążenia estymatorów.

Wynik estymacji zależy od wartości współczynników regularyzacji k i γ . Podczas badań metodą Monte Carlo tak dobrano wartości tych współczynników, aby osiągnąć minimum *a posteriori* błędu średniokwadratowego estymacji. Sytuacja taka obrazuje optymalne właściwości badanych algorytmów. Błąd średniokwadratowy zdefiniowany jest jako wartość oczekiwana kwadratu odległości między wektorem estymat parametrów i ich wartościami prawdziwymi w L^2 [1], [3], [7]:

$$D_{MSE} = E \left\| \hat{\boldsymbol{\theta}}_R - \boldsymbol{\theta}_0 \right\|_2^2 = \text{tr} \mathbf{S}_R + \mathbf{b}_R^T \mathbf{b}_R. \quad (20)$$

Pierwszy składnik powyższej sumy (śląd macierzy kowariancji) charakteryzuje błąd losowy estymacji, a drugi – wynikający z obciążenia błąd systematyczny.

4. Wyniki symulacji i ich analiza

Wyniki analizy właściwości omówionych metod estymacji parametrów przedstawionego modelu nieliniowego zebrano w Tab. 1. W trakcie symulacji metodą Monte Carlo określono *a posteriori* optymalne wartości współczynników regularyzacji (poza metodą GN, która nie wykorzystuje regularyzacji) minimalizujące błąd średniokwadratowy estymacji.

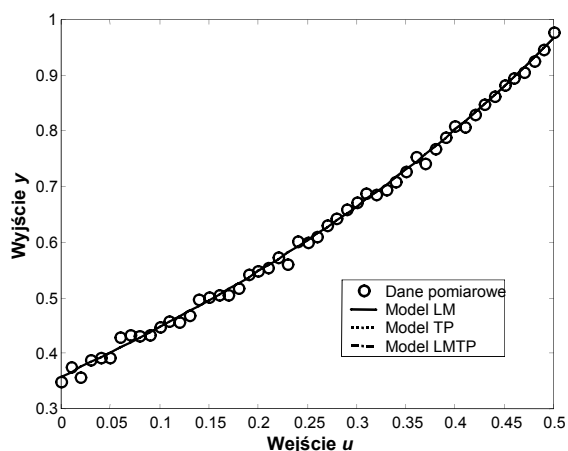
Spośród czterech analizowanych metod najbardziej problematyczny okazał się algorytm Gaussa-Newtona zalecany do estymacji parametrów typowych modeli nieliniowych. Już przy 5 składnikach wchodzących w skład modelu (a tym samym 5 parametrach) nie był on zbieżny, a tym samym nie mógł być uwzględniony w pełnych badaniach przeprowadzonych na modelu z 10 parametrami. Oszacowano jedynie liczbę uwarunkowania numerycznego odwracanej w tej metodzie macierzy i rząd śladu macierzy kowariancji, które wynoszą odpowiednio $3 \cdot 10^{19}$ i 10^{14} . Liczby te wyjaśniają, dlaczego algorytm GN nie radził sobie z estymacją 10 parametrów na podstawie zaszumionych danych pomiarowych.

Tab. 1. Właściwości metod estymacji parametrów modelu nieliniowego.
Tab. 1. Properties of the methods for parameter estimation of the nonlinear model.

Metoda estymacji	GN	LM	TP	LMTP
Współczynnik regularyzacji	—	$k=3 \cdot 10^{-4}$	$\gamma=2 \cdot 10^{-3}$	$k=2 \cdot 10^{-4}$, $\gamma=1.5 \cdot 10^{-3}$
Liczba uwarunkowania numerycznego	$3 \cdot 10^{19}$	$7 \cdot 10^4$	$3 \cdot 10^4$	$3 \cdot 10^4$
Błąd systematyczny ($\mathbf{b}^T \mathbf{b}$)	—	$3.7 \cdot 10^{-1}$	$4.2 \cdot 10^{-2}$	$4.9 \cdot 10^{-2}$
Błąd losowy (trS)	10^{14}	$3.7 \cdot 10^{-1}$	$4.2 \cdot 10^{-2}$	$4.7 \cdot 10^{-2}$

Pozostałe trzy metody, LM oraz zaproponowane w referacie nieliniowa wersja TP i połączenie LMTP, posiadają wbudowany mechanizm poprawiający uwarunkowanie numeryczne estymacji, co przełożyło się na znaczną redukcję liczby uwarunkowania odwracanych macierzy (do rzędu 10^4) i zmniejszenie śladu macierzy kowariancji (do rzędu 10^{-1} dla LM i 10^{-2} dla pozostałych dwóch). Regularyzacja wprowadziła jednak obciążenia estymatorów, których miary (dla przyjętych w symulacjach wartości współczynników regularyzacji) są tego samego rzędu, co ślady macierzy kowariancji. Wynika stąd, że błąd średniokwadratowy estymacji metodą TP był o rząd mniejszy niż metodą LM, co jest równoznaczne z większą dokładnością

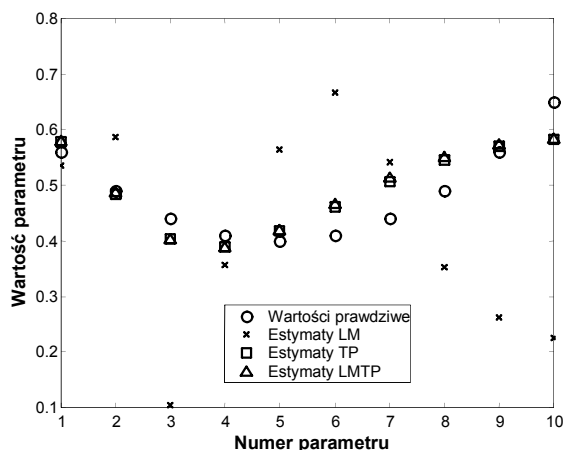
oszacowania parametrów modelu uzyskiwanej metodą TP. W przypadku metody LM wynik estymacji zależy dodatkowo od doboru startowego wektora parametrów, co jest kolejną wadą tego podejścia. Połączona metoda LMTP nie przyniosła w kontekście dokładności pomiaru pośredniego żadnych istotnych korzyści.



Rys. 1. Model nieliniowy dopasowany metodami LM, TP i LMTP do syntetycznych danych pomiarowych.

Fig. 1. Nonlinear model fitted to synthetic measurement data by the LM, TP and LMTP methods.

Na Rys. 1 pokazano dopasowanie modelu nieliniowego do syntetycznych danych pomiarowych uzyskane trzema zregulowanymi metodami estymacji. Jakość dopasowania określona przez błąd średniokwadratowy rezyduów była w tych przypadkach prawie identyczna ($3.5 \cdot 10^{-3}$), co pokazuje, że nie może być ona kryterium oceny dokładności pomiaru pośredniego.



Rys. 2. Przykładowe wyniki estymacji parametrów modelu nieliniowego metodami LM, TP i LMTP.

Fig. 2. Exemplary results of estimation of nonlinear model parameters by the LM, TP and LMTP methods.

Właściwości estymacji poszczególnych parametrów trzema analizowanymi algorytmami zobrazowane są na Rys. 2. Widoczne jest o wiele lepsze oszacowanie uzyskane metodami TP i LMTP w porównaniu z LM, co w tych przypadkach wynika z mniejszego o rząd błędu średniokwadratowego estymacji. Warto zauważyć, że w metodzie LM estymaty słabo określonych parametrów o wysokich numerach pozostały w pobliżu wektora startowego (o wartościach 0.2), co zostało zrekompensowane przez zawyżenie estymat „środkowych” związanych z podobnymi przebiegami wektorów wrażliwości. Odwrotnie, mechanizm zawarty w regularyzacji Tichonowa-Phillipsa powodował „przyciąganie” estymat parametrów słabo określonych w pobliże estymat sąsiednich, co zaowocowało o wiele mniejszym błędem systematycznym (a także losowym – patrz Tab. 1).

5. Podsumowanie i wnioski

Celem referatu było przedstawienie modyfikacji regularyzacji Tichonowa-Phillipsa przystosowującej ją do estymacji parametrów modeli nieliniowych oraz analiza właściwości tej metody. Spodziewano się, że regularyzacja taka będzie dawać dobre wyniki w przypadku źle uwarunkowanej numerycznie estymacji parametrów modeli fizyko-matematycznych charakteryzujących się trudną do redukcji strukturą oraz „regularnym” rozkładem wartości elementów wektora parametrów. Przykładem obiektów o tego typu właściwościach są np. systemy o parametrach rozłożonych opisywane modelami o parametrach skupionych.

Właściwości nieliniowej wersji metody Tichonowa-Phillipsa porównano ze standardowymi algorytmami Gaussa-Newtona i Levenberga-Marquardta oraz zaproponowano metodę łączącą regularyzację LM i TP. Badania przeprowadzone zostały za pomocą metody Monte Carlo. Ich wyniki pokazują, że w przypadku modeli nieliniowych zawierających parametry słabo określone przez dane pomiarowe i jednocześnie charakteryzujące się „regularnym” rozkładem wartości w wektorze parametrów, najlepsze wyniki daje metoda Tichonowa-Phillipsa. Wynika to w pierwszym rzędzie z postaci mechanizmu regularyzacji, który uwzględnia aprioryczną wiedzę o parametrach. Metoda Gaussa-Newtona nie była zbieżna w analizowanym przypadku z powodu złego uwarunkowania odwracanej macierzy. Połączona metoda LMTP nie wniosła w porównaniu z TP istotnych korzyści.

Podsumowując, w referacie pokazano, iż w pewnych sytuacjach zaproponowana nieliniowa wersja regularyzacji identyfikacji parametrycznej metodą Tichonowa-Phillipsa posiada lepsze, w porównaniu z przeanalizowanymi metodami, właściwości metrologiczne w pomiarach pośrednich. Dotyczy to analizowanej klasy modeli, tzn. modeli fizyko-matematycznych o trudnej do redukcji strukturze i regularnym rozkładzie wartości w wektorze parametrów. Uzyskane wyniki obrazują znany mechanizm poprawy dokładności pomiarów będącej efektem uwzględnienia apriorycznej wiedzy o badanym obiekcie.

6. Literatura

- [1] Gajda J., Szyper M.: Modelowanie i badania symulacyjne systemów pomiarowych. Firma Jartek s.c., Kraków 1998.
- [2] Hoerl A.E.: Application of ridge analysis to regression problems. Chem. Eng. Progress 58, 1962, 54-59.
- [3] Hoerl A.E., Kennard R.W.: Ridge regression: biased estimation for non-orthogonal problems. Technometrics 12, 1970, 55-62.
- [4] Johansen T.A.: On Tikhonov regularization, bias and variance in nonlinear system identification. Automatica 33, 1997, 441-446.
- [5] Levenberg K.: A method for the solution of certain nonlinear problems in least squares. Quart. Appl. Math. 2, 1944, 164-168.
- [6] Marquardt D.W.: An algorithm for least squares estimation of nonlinear parameters. J. SIAM 11, 1963, 431-441.
- [7] Marquardt D.W.: Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation. Technometrics 12, 1970, 591-612.
- [8] Morawski R.Z., Szczeciński L., Barwicz A.: Deconvolution algorithms for instrumental applications: a comparative study. J. Chemometrics 9, 1995, 3-20.
- [9] Mroczka J.: Metrologia w procesie poznania. W: Barzykowski J. (red.): Współczesna metrologia – zagadnienia wybrane. WNT, Warszawa 2004.
- [10] Phillips D.L.: A technique for the numerical solution of certain integral equations of the first kind. J. ACM 9, 1962, 84-97.
- [11] Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P.: Numerical Recipes in C: the Art of Scientific Computing. Cambridge University Press, Cambridge, MA, 1992.
- [12] Tikhonov A.N., Arsenin V.Y.: Solutions of Ill-Posed Problems. Winston, Washington, D.C., 1977.
- [13] Twomey S.: On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature. J. ACM 10, 1963, 97-101.