

# Ćwiczenie 2. Identyfikacja parametrów funkcji statycznych — regresja liniowa i wielomianowa

Dominik Sierociuk  
Politechnika Warszawska  
Instytut Sterowania i Elektroniki Przemysłowej

OKNO, Modelowanie i Identyfikacja Systemów Dynamicznych

## 1 Wprowadzenie

W ćwiczeniu tym zostanie przedstawiona metoda identyfikacji parametrów funkcji statycznych bazującej na metodzie najmniejszych kwadratów (ang. Least-Squares (LS) [Rozdział 3 skryptu]). Zostaną przedstawione algorytmy regresji liniowej oraz wielomianowej.

Jeżeli chcemy wyznaczyć parametry funkcji  $y = ax + b$ , która przechodzi przez dwa punkty,  $(x_1, y_1)$  i  $(x_2, y_2)$  możemy zapisać to w postaci układu dwóch równań:

$$\begin{aligned}y_1 &= ax_1 + b \\ y_2 &= ax_2 + b\end{aligned}$$

Zapisując ten układ równań w postaci macierzowej otrzymamy

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

co możemy zapisać w skrócie:

$$b = Ax$$

Układ równań możemy rozwiązać macierzowo, jeżeli macierz  $A$  jest odwracalną. Będzie miało miejsce, gdy będą to dwa różne punkty. W innym przypadku – dwóch tych samych punktów – macierz będzie miała dwa wiersze liniowo zależne i układ równań będzie miał nieskończenie wiele rozwiązań (przez dwa te same punkty możemy przeciągnąć nieskończenie wiele prostych). Jeżeli więc macierz  $A$  będzie odwracalna, to jednoznaczne rozwiązanie będzie dane następująco:

$$x = A^{-1}b$$

To wszystko jest takie proste, jeżeli mamy do czynienia z tak trywialnym przykładem, rzeczywistość jest jednak znacznie bardziej złożona. Zazwyczaj mamy

do czynienia z sytuacją, w której będziemy dysponowali znacznie większą liczbą punktów niż szukanych parametrów (czyli będziemy mieć do czynienia z nadokreślonym układem równań), co więcej punkty te będą obciążone zakłóceniami. O ile nadmiarowa liczba równań nie byłaby wielkim problemem przy idealnych pomiarach, bo po prostu nadmiarowe równania były by liniowo zależne (więc nie wносиły by żadnej dodatkowej informacji) i można by je zignorować. Natomiast w przypadku danych, które nie są idealne, takiego rozwiązania nie możemy zrobić. Zakłócenia powodują, że mamy do czynienia z układem równań sprzecznych (Niestety, podstawowe twierdzenie geometrii wykreślnej, o tym, że przez dowolne trzy punkty można przeprowadzić jedną prostą, o ile ta prosta będzie dostatecznie gruba, tutaj nie zadziała). Będziemy potrzebowali więc narzędzia, które będzie w stanie wyznaczyć rozwiązanie, będące najbliższemu poszukiwanemu. Takim narzędziem jest pseudoodwrotność macierzy Moore'a-Penrose'a oznaczanej jako  $A^\dagger$ . Operacja ta, przy rozwiązywaniu układów równań sprzecznych, daje wynik będący minimalizacją odległości (kwadratu błędu) pomiędzy otrzymaną funkcją, a danymi punktami. Przyjmując oznaczenia:

$$\theta = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \varphi_i^T = [x_i \ 1], \quad \Phi = \begin{bmatrix} \varphi_1^T \\ \vdots \\ \varphi_N^T \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$$

Zadanie rozwiązywane przez pseudoodwrotność możemy więc zdefiniować jako minimalizację następującej funkcji:

$$V(\theta) = \sum_{i=1}^N (\varphi_i^T \theta - y_i)^2$$

czyli kwadratu pomiędzy wartością estymowanej funkcji dla danych szukanych parametrów i wartością z pomiarów, a zadanie minimalizacji możemy zapisać teraz jako:

$$\hat{\theta} = \arg \min_{\theta} V(\theta)$$

Operacja ta podaje więc najlepszy, wg. tego sensu, wynik, jaki możemy otrzymać z tych zakłóconych danych.

Pseudoodwrotność możemy wyznaczyć z następujących zależności:

$$A^\dagger = \begin{cases} (A^T A)^{-1} A^T, & \text{dla } \text{rank}(A) = m \\ A^T (A A^T)^{-1}, & \text{dla } \text{rank}(A) = n. \end{cases}$$

Rozwiązanie układu równań jest wyznaczane więc następująco

$$x = A^\dagger b$$

Szczegóły wyznaczania pseudoodwrotności i rozwiązywania z jej użyciem układów równań w skrypcie.

Innym sposobem wyznaczenia macierzy pseudoodwrotnej jest skorzystanie z rozkładu macierzy względem wartości szczególnych (SVD). Jeżeli rozkład SVD

macierzy  $A$  dany jest następująco  $A = U\Sigma V^T$  to macierz pseudoodwrotną możemy wyznaczyć następująco:

$$A^\dagger = V\Sigma^\dagger U^T. \quad (1)$$

Jako, że macierz  $\Sigma$  zawiera elementy niezerowe tylko na jej przekątnej jej pseudoodwrotność będzie oznaczała tylko inversję tych niezerowych wartości.

$$\Sigma^\dagger = \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix}, \quad (2)$$

gdzie

$$D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & d_N \end{bmatrix}, D^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \dots & 0 \\ 0 & \frac{1}{d_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \frac{1}{d_N} \end{bmatrix}, \quad (3)$$

a  $d_i$  są wartościami szczególnymi macierzy  $A$ . Czyli, gdy mamy dokonany rozkład SVD macierzy  $A$  pseudoodwrotność macierzy możemy wykonać przez prostą inversję wartości na diagonalu macierzy  $\Sigma$  oraz dokonanie dwóch mnożeń macierzy.

Co więcej, rozkład SVD umożliwia nie tylko wyznaczenie pseudoodwrotności, ale daje dużo użytecznych informacji z punktu widzenia rozwiązywania układu równań. Przede wszystkim, liczba niezerowych wartości szczególnych macierzy mówi nam o rzędzie tej macierzy, czyli jeżeli po dokonaniu rozkładu SVD otrzymamy choć jedną wartość szczególną zerową, oznaczać to będzie, że macierz nie jest pełnego rzędu, czyli, że mamy choć jedną kolumnę (bądź wiersz gdy mniej wierszy niż kolumn) liniowo zależną od innej kolumny, co będzie oznaczało niemożność otrzymania jednoznacznego rozwiązania. Macierze  $U$  i  $V$  zawierają im mniej bazy wektorowe odpowiednich rzutów geometrycznych, których zastosowanie umożliwia takie przekształcenie danych, aby ich klasyfikacja (rozróżnienie) było jak najefektywniejsze, co ma decydujące znaczenie na przykład w metodzie Principal Component Analysis (PCA).

## 2 Przykłady regresji liniowej

**Przykład 1.** Jako przykład weźmy funkcję liniową  $y = 0.5x + 0.6$ , której wartości zostały zmierzone z zakłóceniami, zakładając nieznaną wartość jej parametrów zidentyfikujemy je. Macierze będą miały przywołaną wcześniej postać

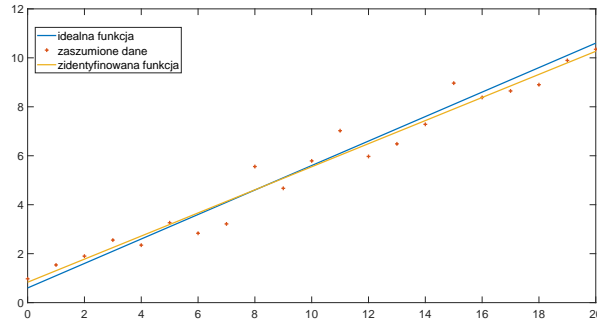
$$\theta = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \varphi_i^T = [x_i \ 1], \quad \Phi = \begin{bmatrix} \varphi_1^T \\ \vdots \\ \varphi_N^T \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix},$$

gdzie  $x_i$  są kolejnymi wartościami argumentów funkcji, a  $y_i$  kolejnymi zmierzonymi wartościami tej funkcji. Rozwiązanie tego układu równań zostało otrzymane

z użyciem pseudoodwrotności (poprzez komendę `pinv` w Matlabie),

$$x = A^\dagger b.$$

Otrzymane wartości to  $\hat{a} = 0.4716$  i  $\hat{b} = 0.8338$ , a wyniki w formie graficznej przedstawione są na Rysunku 1. Jak widać parametry wyestymowane różnią się od idealnych (szczegółności parametr  $b$ ), natomiast suma kwadratu błędu pomiędzy pomiarami a wyestymowaną prostą wyniosła 5.8791, natomiast błędu pomiędzy pomiarami a idealną prostą wyniosła 6.5517. Algorytm pseudoodwrotności określił więc parametry, które dawały mniejszy błąd niż te, dla których eksperyment przeprowadzono, co ma związek tylko z danymi zakłóceniami.



Rysunek 1: Wyniki identyfikacji funkcji liniowej

**Przykład 2.** Przykład identyfikacji parametrów układu ciągłego.

Metodę regresji liniowej możemy również zastosować do identyfikacji parametrów układu ciągłego poprzez taki zapis jego odpowiedzi analitycznej, aby miała ona postać równania liniowego (oczywiście taką modyfikację zapisu równania analitycznej odpowiedzi, aby miała postać równania liniowego, możemy wykonać tylko w niektórych przypadkach).

Weźmy układ inercyjny o transmitancji

$$G(s) = \frac{1}{as + 1} \quad (4)$$

jego analityczna odpowiedź skokowa wyznaczona została w następujący sposób

$$y(t) = \mathcal{L}^{-1}\left\{\frac{1}{as + 1} \frac{1}{s}\right\} = 1(t) - e^{-at}$$

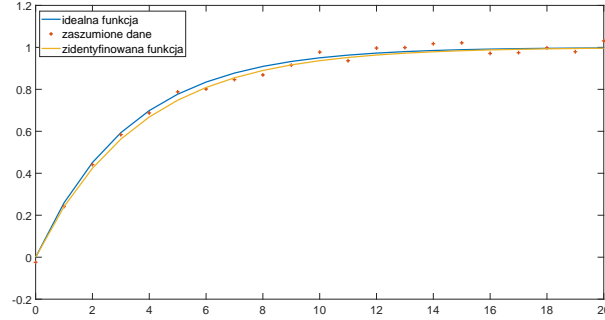
przekształcając i logarytmując obie strony równania otrzymujemy

$$\ln(1 - y(t)) = -at$$

co możemy zapisać w postaci macierzowej dla kolejnych próbek

$$\theta = \begin{bmatrix} a \end{bmatrix}, \quad \varphi_i^T = [-t_i], \quad \Phi = \begin{bmatrix} \varphi_1^T \\ \vdots \\ \varphi_N^T \end{bmatrix}, \quad Y = \begin{bmatrix} \ln(1 - y(t_1)) \\ \vdots \\ \ln(1 - y(t_N)) \end{bmatrix}$$

gdzie  $t_i$  są kolejnymi wartościami czasu dla poszczególnych próbek.



Rysunek 2: Wyniki identyfikacji układu inercyjnego

Wyniki identyfikacji dla parametru modelu  $a = 0.3$  zostały przedstawione na Rysunku 2, a estymowany parametr ma wartość  $\hat{a} = 0.2758$ . Otrzymana wartość sumy kwadratu błędu pomiędzy pomiarami a wyestymowaną odpowiedzią wyniosła 0.0125, natomiast błędu pomiędzy pomiarami a idealną odpowiedzią wyniosła 0.0121. Powodem, dla którego błąd estymaty jest większy niż idealnego układu, jest pojawiająca się osobliwość przy wyznaczaniu  $\ln(1 - y(t))$ . Osobliwość ta powoduje otrzymanie, przy wartości odpowiedzi powyżej 1 spowodowanej szumem pomiarowym, wartości zespolonej. Dlatego z otrzymywanej wartości logarytmu naturalnego zostaje wybierana tylko wartość rzeczywista. Konieczność takiej modyfikacji implikuje większy błąd niż dla układu idealnego. Natomiast otrzymane wyniki (zgodność odpowiedzi) pozostały na zadowalającym poziomie.

### 3 Przykład regresji wielomianowej

Metodykę stosowaną przy wyznaczaniu parametrów równań stopnia pierwszego możemy również zastosować dla wielomianów dowolnego rzędu.

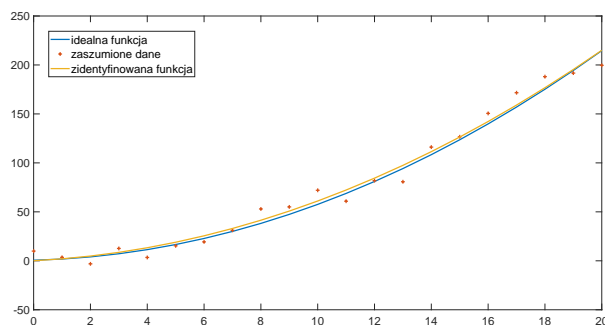
Za przykład weźmy funkcję  $y = 0.5x^2 + 0.7x + 0.6$ , której wartości zostały zmierzone z zakłóceniami.

Zakładamy więc identyfikowaną postać funkcji  $y = ax^2 + bx + c$ . Macierze dla procesu identyfikacji będą miały następującą postać

$$\theta = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad \varphi_i^T = [x_i^2 \ x_i \ 1], \quad \Phi = \begin{bmatrix} \varphi_1^T \\ \vdots \\ \varphi_N^T \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}.$$

Otrzymane wartości to  $\hat{a} = 0.4645$ ,  $\hat{b} = 1.4532$  i  $\hat{c} = 0.0764$ , a wyniki w formie graficznej przedstawione są na Rysunku 3. Jak widać pomimo dość znacznych różnic parametrów estymowanych i rzeczywistych, estymowana funkcja jest

bardzo bliska idealnej. Co więcej, otrzymana wartość sumy kwadratu błędu pomiędzy pomiarami a wyestymowaną parabolę wyniosła  $1.6463e + 03$ , natomiast błędu pomiędzy pomiarami a idealną parabolą wyniosła  $1.7737e + 03$ . Co pokazuje efektywność użytego algorytmu i wpływ szumu pomiarowego na wynik.



Rysunek 3: Wyniki identyfikacji paraboli

## 4 Plan ćwiczenia

1) Przeprowadzić identyfikację dla przynajmniej trzech równań prostej, trzech parabol i trzech równań trzeciego rzędu. Przeanalizować wpływ poziomu zakłóceń i ilości danych (liczby punktów czy wielkości przyrostu argumentu) na otrzymany wynik estymacji oraz porównanie błędu estymacji i pomiarowego.

2) Kondensator naładowany jest do pewnego napięcia  $U_0$ , po czym rozładowuje się poprzez podłączony w chwili  $t = 0$  rezystor. Napięcie na kondensatorze  $U_c$  jest mierzone z dokładnością do 5V w kolejnych momentach. Wyniki pomiarów w tabeli

t [s]	$U_c$ [V]
0	100
1	75
2	55
3	40
4	30
5	20
6	15
7	10

Wiadomo, że postać analityczna odpowiedzi dana jest następująco:

$$U_c(t) = U_0 e^{-at}$$

Przeprowadzić identyfikację parametrów  $U_0$  i  $a$ .

3)W wyniku badań w tunelu aerodynamicznym otrzymano dla modelu samolotu dane o zależności kąta odchylenia steru  $\delta$  gwarantującego równowagę siły nośnej i siły ciężenia (samolot będzie leciał trajektorią prostoliniową, horyzontalną) przy danej prędkości  $v$  strumienia powietrza. Wyniki pomiarów przedstawia poniższa tabela, przy czym przez  $n$  oznaczono liczbę pomiarów przy danej prędkości.

$v$ [m/s]	$\delta$ [°]	$n$
80	-3° 44'	8
90	-2° 58'	12
100	-2° 16'	11
110	-1° 39'	9
120	-1° 21'	14
140	-0° 38'	6
160	-0° 07'	9
180	0° 10'	12
200	0° 35'	10

Wiadomo, że krzywa balansowania określona jest zależnością

$$\delta = a_0 + \frac{a_1}{v^2}$$

Należy zidentyfikować parametry  $a_0$  i  $a_1$ , pamiętając o przeliczeniu minut kątowych na stopnie oraz o uwzględnieniu wielokrotności danych pomiarów.

#### 4.1 Sprawozdanie

W sprawozdaniu z ćwiczenia przedstawić przede wszystkim wnioski jakie nasunęły się podczas przeprowadzania ćwiczenia, można je zilustrować szeregiem otrzymanych wyników. Należy przedstawić także rozwiązanie ćwiczeń nr 2 i 3.