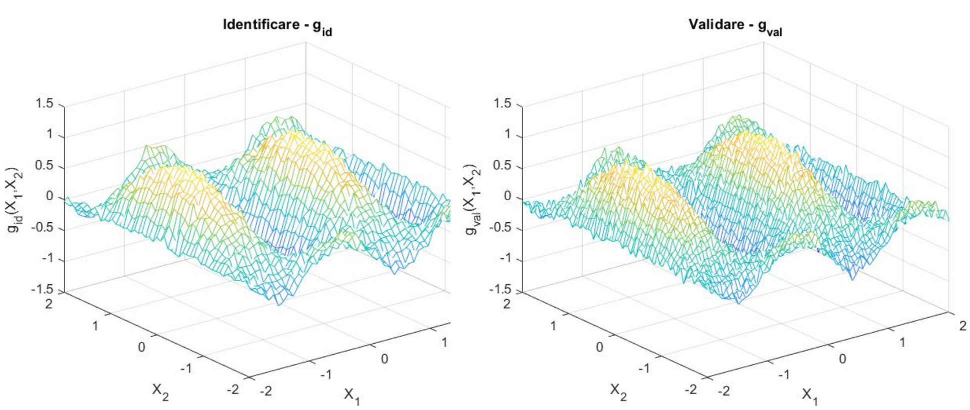
Raport

Zigler Alexandru

November 2021

Introducere

Considerăm o funcție $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$, dată printr-un set de date de identificare ce conține intrările și ieșirea funcției sub forma unor eșantioane. Acest lucru se poate observa în Fig. 1 (a). Folosind metoda regresiei liniare, realizăm un aproximator polinomial de grad variabil pentru a obține un model al funcției. Acest model se validează ulterior folosind un alt set de date reprezentat în Fig. 1 (b). Dorim să găsim gradul pentru care eroarea medie pătratică a modelului este minimă în raport cu setul de validare, având în vedere și conceptul de supraantrenare.



(a) Identificare (b) Validare

Figure 1

Structura aproximatorului

Pentru ca gradul m să fie variabil, determinăm formula numărului de termeni al aproximatorului, pe care îl notăm cu $u(m)$.

$$
\begin{aligned}
&u(m) = u(m-1) + (m+1) \\
&u(m-1) = u(m-2) + m \\
&u(m-2) = u(m-3) + (m-1) \\
&\qquad\qquad\vdots \\
&u(3) = u(2) + 4 \\
&u(2) = u(1) + 3 \\
&u(1) = 3 \\
\hline
&u(m) = 3 + 3 + 4 + \cdots + m + (m+1) \\
&u(m) = \tfrac{(m+1)(m+2)}{2}
\end{aligned}
$$

Scriem $g(x_1, x_2) = \theta_1 \cdot 1 + \theta_2 \cdot x_1 + \theta_3 \cdot x_2 + \theta_4 \cdot x_1^2 + \theta_5 \cdot x_1 x_2 + \theta_6 \cdot x_2^2 + \dots$ Ecuația are $u(m)$ termeni și se poate scrie și sub formă matriceală:

$$
\begin{bmatrix}
1 & x_{11} & x_{21} & x_{11}^2 & x_{11}x_{21} & x_{21}^2 & \dots \\
1 & x_{11} & x_{22} & x_{11}^2 & x_{11}x_{22} & x_{22}^2 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\
1 & x_{1n} & x_{2n} & x_{1n}^2 & x_{1n}x_{2n} & x_{2n}^2 & \dots
\end{bmatrix}
\begin{bmatrix}
\theta_1 \\ \theta_2 \\ \vdots \\ \theta_{u(m)}
\end{bmatrix}
=
\begin{bmatrix}
y_{11} \\ y_{12} \\ \vdots \\ y_{21} \\ \vdots \\ y_{nn}
\end{bmatrix}
$$

Construim matricea de regresori $\Phi \in \mathcal{M}_{n^2, u(m)}$, unde n este dimensiunea celor 2 vectori de intrare X_1 și X_2.

Pentru m=1, linia se construiește din elementele $[1, x_1, x_2]$.

Pentru $m \geq 2$, se rețin elementele generate la ordinul anterior și se adaugă elementele corespunzătoare ordinului curent. Acestea se obțin din combinații de puteri ale lui $x_1 x_2$ (i.e. $x_1^{m-p} x_2^p$, $p = \overline{0, \mathrm{m}}$).

Se rezolvă sistemul $\Phi \cdot \Theta = Y$, obținându-se astfel setul de parametri Θ.

$\hat{Y} = \Phi \cdot \Theta$

Implementare MATLAB

Structură:

- script principal: ”proiect”
- funcția ”generateLine”

Amintim câteva detalii semnificative de implementare.

Se selectează x_{1i} și x_{2j} din vectorii de intrare pentru fiecare linie din matricea Φ, în funcție de $k = \overline{0, \mathrm{n}^2}$.

Funcția "generateLine" preia valorile menționate mai sus și gradul curent m, furnizând pe baza lor un vector care se va folosi în construcția matricei linie cu linie.

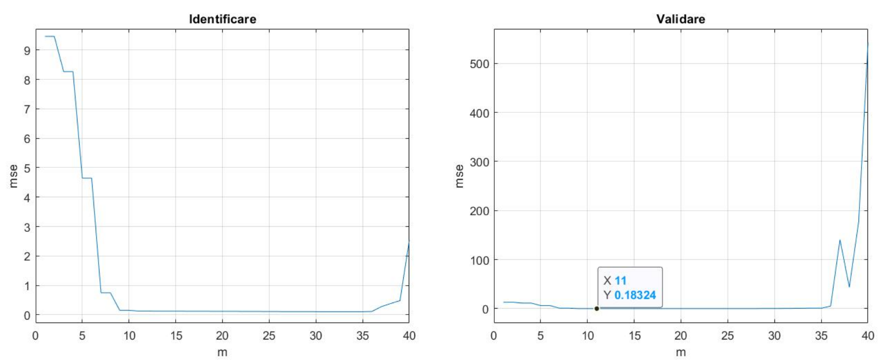
Ecuația $\Phi \cdot \Theta = Y$ se rezolvă folosind operatorul "\". Rezultatul aproximării (i.e. $\Phi \cdot \Theta$) este de forma unui vector coloană de lungime n^2. Folosind funcția "reshape", acesta se rearanjează sub forma unei matrice $\hat{Y} \in \mathcal{M}_{n,n}$.

Se calculează în mod analog matricea Φ corespunzătoare setului de validare, apoi ieșirea aproximată folosind vectorul de parametri Θ, obținut la pasul de identificare.

Se analizează performanța aproximării, calculând eroarea medie pătratică și se repetă procesul pentru fiecare m din intervalul considerat.

Rezultate de acordare

Se calculează eroarea medie pătratică (mse) pentru fiecare grad (m) din intervalul considerat atât pentru setul de identificare, cât și pentru cel de validare. Aceste date se păstrează pentru a compune grafice care ilustrează evoluția erorilor în funcție de grad - Fig. 2.



(a) (b)

Figure 2: Evoluția erorilor medii pătratice

Observăm că inițial ($m = \overline{1, 7}$) erorile sunt relativ mari(mse $\in (0.7, 14)$) și încep să scadă pe măsură ce crește gradul. În cazul setului de validare, valoarea minimă a erorii este 0.183 pentru m=11. Pentru $m \geq 36$, erorile cresc și constatăm că ajungem la supraantrnare.

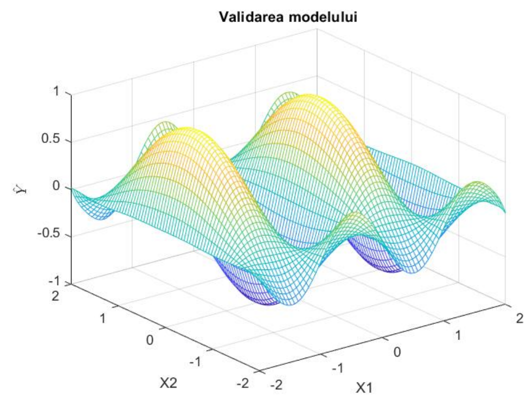


Figure 3: Aproximarea modelului pentru mse minim

Odată ce cunoaștem gradul optim m=11 pentru care aproximarea este cea mai bună, recalculăm matricea $\hat{Y}$ și o reprezentăm grafic în Fig. 3.

Comparăm modelul aproximat $\hat{g}_{val}(X_1, X_2)$ cu funcția dată $g_{val}(X_1, X_2)$ din Fig. 1(b). Observăm că $\hat{g}_{val}(X_1, X_2)$ este mai netedă decât $g_{val}(X_1, X_2)$, care este perturbată de zgomot de medie zero.

În concluzie, cu ajutorul metodei regresiei liniare, am realizat un aproximator polinomial cu grad variabil. Conceptul de eroare medie pătratică a fost util în determinarea performanțelor fiecărui grad. Astfel, am găsit cel mai apropiat model de sistemul considerat inițial.

smcasmţ