UNIVERSITATEA TEHNICĂ
DIN CLUJ-NAPOCA

Proiect Identificarea Sistemelor-partea I

Modelarea unei funcții necunoscute

Noiembrie 2023

Nume studenți:
Florean Ioana-Valentina
Tăbîrcea Georgiana-Maria
Ciceu Robert

Setul de date :proj_fit_14

Grupa:30135
Semigrupa:1

Cuprins

Descrierea problemei

Această prezentare se concentrează pe dezvoltarea unui aproximat polinomial pentru o funcție necunoscută,statică și neliniară, folosind date de intrare-ieșire furnizate.Scopul este să devoltăm un model care să poată estima funcția dată, în ciuda narurii sale neliniare. Ieșirea funcției este afectată de zgomot aditiv de tip Gaussian, cu o medie zero.

Acest proces de modelare se bazează pe doua seturi distincte de date:

- Setul de identificare , notat cu "id", folosind pentru antrenarea modelului.
- Setul de validare, notat cu "val", utilizat pentru evaluarea performantei modelului dezvoltat.

Descrierea structurii aproximatorului și a procedurii de găsire a parametrilor

Aproximatorul polinomial trebuie să aibă gradul configurabil, iar antrenarea modelului constă în găsirea vectorului optim de parametrii pentru care polinomul se apropie cel mai mult de funcția necunoscută, dată de setul de date de identificare.

Structura aproximatorului polinomial:

Forma generala:

$y(k)= [\phi1(k), \phi2(k). . . , \phi n(k)] \cdot [\theta1, \theta2, . . . , \theta n]$ ^T + e(k)

Unde:

- $\phi1(k), \phi2(k),..., \phi n(k)$ sunt regresorii
- $\Theta1, \theta2, . . . , \theta n$ sunt vectori necunoscuți
- e(k) este eroarea modelului

Procedura de găsire a parametrilor:

Antrenarea modelului constă în găsirea vectorului optim de parametri θ pentru care polinomul y(k) se apropie cât mai bine de funcția necunoscută. Aceasta se realizează prin minimizarea erorii pătratice medii (MSE) între ieșirile estimate și cele reale ale funcției.

În continuare vă vom prezenta codul care stă la baza regresiei liniare având comentarii explicite.

```
load("proj_fit_14.mat")
id.X;
id.Y;
val.X;
val.Y;
figure;
mesh(id.X{1},id.X{2},id.Y)
mesh(val.X{1},val.X{2},val.Y)
title("Grafic Initial")
phi = [];
phi_val=[];
theta=[];
n=30;
MSE1=zeros(1,n);
MSE2=zeros(1,n);
```

```
for n=1:30
%matricea phi de identificare
N=length(id.X{1}); %nr de puncte din setul de identificare
il = 1; %indexul liniei
for i = 1:N
for j = 1:N
ic = 1; %indexul coloanei
for m = 0:n %putere
for w = 0:n %putere
if m + w <= n
if m == 0 && w == 0
phi(il, ic) = 1; % Prima coloană este întotdeauna 1
else
phi(il, ic) = id.X{1}(i)^m * id.X{2}(j)^w; % Calculează valorile pentru celelalte coloane
end
ic = ic + 1;
end
end
end
```

```
il = il + 1;
end
end
% matricea phi de validare
M=length(val.X{1});
il_val = 1;
for z = 1:M
for b= 1:M
ic_val = 1;
for p= 0:n
for t = 0:n
if p+t <= n
if p == 0 && t == 0
phi_val(il_val, ic_val) = 1; % Prima coloană este întotdeauna 1
else
phi_val(il_val, ic_val) = val.X{1}(z)^p * val.X{2}(b)^t; % Calculează valorile pentru celelalte coloane
end
ic_val = ic_val + 1;
end
end
end
il_val = il_val + 1;
end
end
```

```
Y_id=reshape(id.Y,N*N,1);
theta=phi\Y_id;
y_aprox_id=phi*theta;
y_aprox1_id=reshape(y_aprox_id,N,N);
Y_val=reshape(val.Y,M*M,1);
y_aprox_val=phi_val*theta;
y_aprox1_val=reshape(y_aprox_val,M,M);
if n==4
figure
mesh(y_aprox1_id)
title("Aproximare si identificare")
figure
mesh(y_aprox1_val)
title("Aproximare si validare")
end
%MSE pentru identificare
e1=ones(N, 1); %matrice care sa stocheze toate valorile e1
for k=1:N
e1(k)=id.Y(k)-y_aprox1_id(k);
end
```

```
MSE1(n)=1/N*sum(e1.^2);
%MSE pentru validare
e2=ones(M, 1); %matrice care sa stocheze toate valorile e2
for K=1:M
e2(K)=val.Y(K)-y_aprox1_val(K);
end
MSE2(n)=1/M*sum(e2.^2);
[minim,min_gr]=min(MSE2);
end
figure
plot(1:n,MSE1)
title("Eroarea Medie Patratica 1")
figure
plot(1:n,MSE2)
title("Eroarea Medie Patratica 2")
fprintf('MSE MINIM ESTE IN PUNCTUL:%f \nmin grad%d',min(MSE2),min_gr)
```

Caracteristici esențiale ale soluției

Datele de identificare și validare sunt reprezentate de variabilele id.X, id.Y, val.X, și val.Y.

Odată ce datele sunt încărcate , utilizăm funcția "mesh()" pentru a afișa graficele inițiale tridimensionale.

Mai apoi, am construit matricile phi si phi_val, care sunt esentiale pentru procesul de identificare si validare a modelului.

Aceasta constructie presupune generarea si combinarea unor puteri ale datelor de intrare, cu scopul de a crea un set complex de caracteristici utilizate pentru modelare

Am determinat parametrii modelului, reprezentati de vectorul theta, prin rezolvarea unui sistem linear in sensul celor mai mici patrate, realizata prin operatorul \(“backslash()”). Dupa determinarea parametrilor,se efectueaza aproximarea datelor de identificare si validare.

Urmatoarea etapa consta in calcularea erorilor medii patratice(MSE), unde MSE reprezinta o masura a discrepantei intre datele reale si cele approximate de model.

In final, se afiseaza graficele erorilor pentru identificare si validare si se evidentiaza punctul in care MSE pentru validare atinge minimul, oferind astfel o evaluare concludenta a performantei modelului.

Rezultate de acordare

În continuare, vom analiza rezultatele obținute în urma variației numărului de termeni 'n',examinănând eroarea medie pătratică (MSE) îin intervalul de la 1 la 30. Am ales gradul polinomului 30 deoarece asa se observă cel mai bine la eroarea medie pătratică de validare că va crește.

Fig 1: Se obervă o scădere a erorii până în n=4. Pentru n>4 sistemul va avea o eroare medie pătratică constantă.

Fig 2: Se observă o scădere a erorii până în n=4. După eroarea medie pătratică este constantă până în punctul 23 unde observăm o urcare a erorii.

Valoarea minimă a parametrului este n=4, pentru că y_hat se apropie cel mai bine de valorile reale y pe setul de date de identificare.

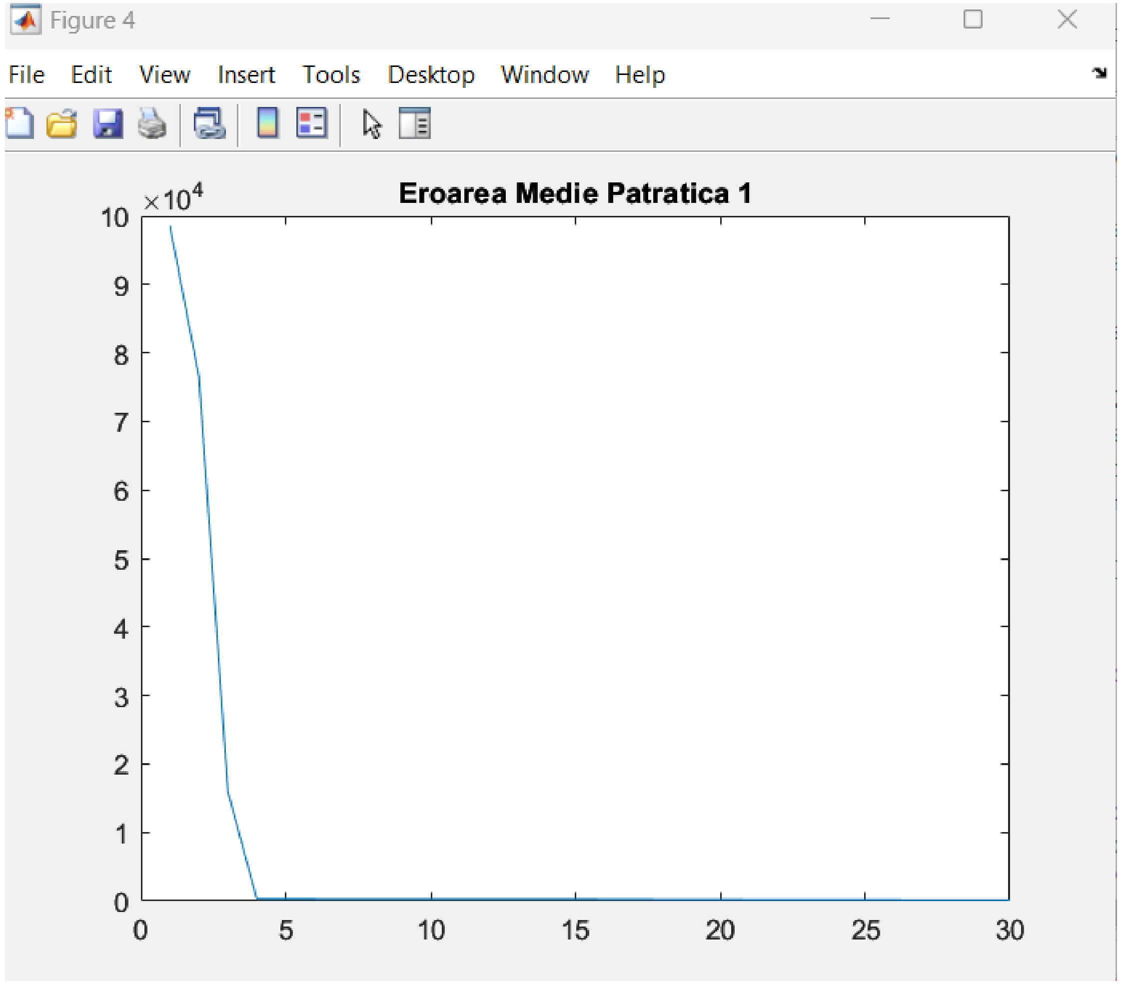


Figure 4
File Edit View Insert Tools Desktop Window Help
Eroarea Medie Patratica 1
$\times 10^4$
10
9
8
7
6
5
4
3
2
1
0
0 5 10 15 20 25 30

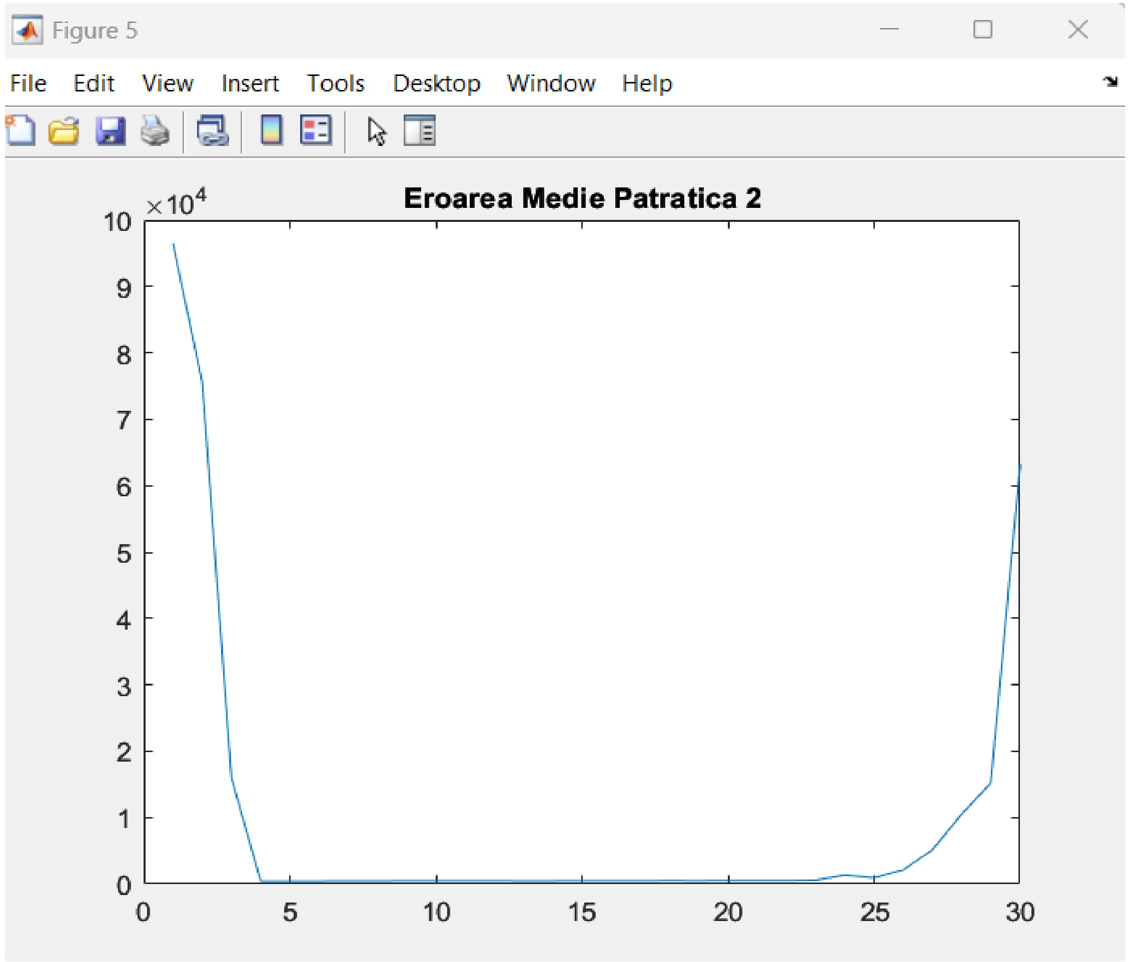
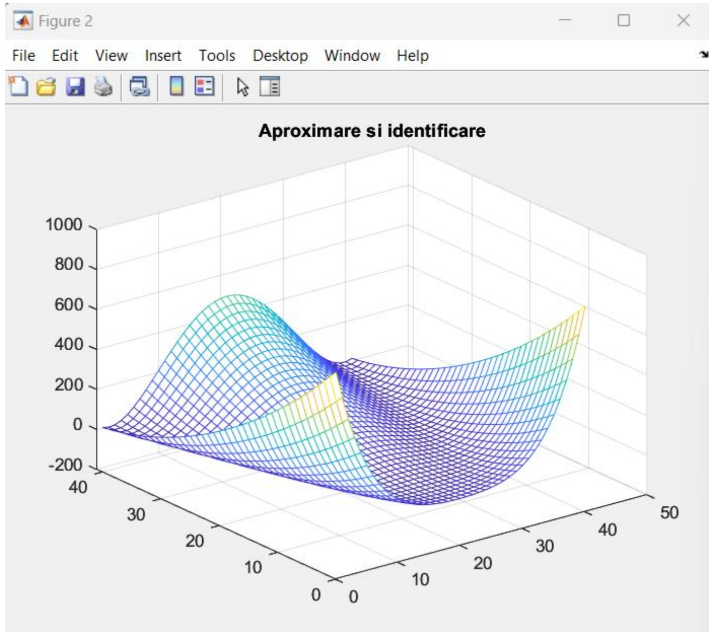
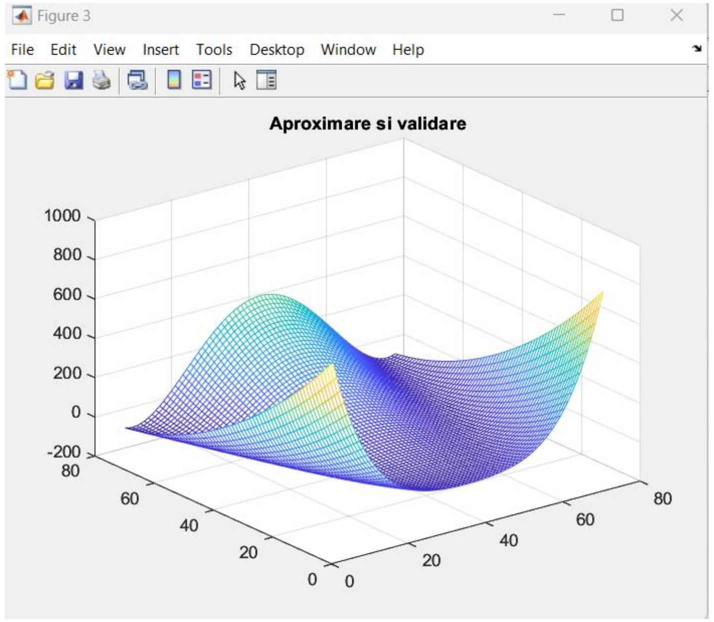


Figure 5
File Edit View Insert Tools Desktop Window Help
Eroarea Medie Patratica 2
$\times 10^4$
10
9
8
7
6
5
4
3
2
1
0
0 5 10 15 20 25 30

Grafice reprezentative

- Fig 3(Date identificare): Acest grafic reflecta procesul de ajustare a modelului la setul de identificare. Pentru valoarea optima a parametrului”n”, in acest caz, n=4, modelul ofera o aproximatie semnificativa si precisa a datelor reale. Acest aspect este evidentiat prin modul in care suprafata graficului se apropie de forma datelor de identificare.
- Fig 4(Date validare): Modelul allege un numar optim de parametric, evitand astfel atat sub-ajustarea,cat si supra-ajustarea si manifestandu-se intr-o generalizare eficienta pe datele de validare.





Concluzie

În concluzie, codul implementează o metodă de identificare a unui sistem utilizând regresia liniară polinomială de gradul n, evaluată prin eroarea medie pătratică (MSE). Un aspect important al analizei este găsirea gradului optim al polinomului, care a fost determinat pe baza minimului local al MSE pe setul de validare.

Sugestie de imbunatatire si optimizarea codului:

Se poate îmbunătăți eficiența codului prin utilizarea unor funcții și vectorizarea operațiilor