

## Laborator 4

Considerăm problema de regresie descrisă la cursul 3 în Slides Lecture 3b: dată fiind o mulțime de exemple de antrenare  $S = \{(x_i, u_i)\}_{i=1,n}$  găsiți polinomului optim care descrie cât mai bine corespondența dintre punctele  $x_i$  și etichetele  $u_i$ .

Fișierul *exempleAntrenare.mat* conține 10 exemple de antrenare, linia  $i$  conținând perechea  $(x_i, u_i)$ . Încărcați conținutul acestui fișier folosind funcția *load*.

Realizați următoarele:

- scrieți funcția ***ploteazaExemple.m*** care plotează exemplele de forma  $(x_i, u_i)$ , cu  $x_i$  reprezentat pe axa Ox și  $u_i$  reprezentat pe axa Oy. Apelați funcția pentru mulțimea de antrenare  $S$ .
- scrieți funcția ***gasestePolinomOptim.m*** care pentru o mulțime de antrenare  $S$  și pentru un grad  $n$  găsește (folosiți funcția Matlab *polyfit.m*) polinomul optim  $P_{S,n}$  de grad  $n$  care minimizează suma erorilor pătratice dintre eticheta prezisă  $P(x_i)$  și eticheta  $u_i$ .
- scrieți funcția ***ploteazaGraficPolinom.m*** care plotează graficul funcției polinomiale asociate polinomului  $P_{S,n}$  pe intervalul  $[0,1]$  (folosiți funcția Matlab *polyval.m*).
- scrieți funcția ***calculeazăEroare.m*** care calculează eroarea  $E$  dintre predicțiile realizate pentru punctele  $x_i$  și etichetele adevărate  $u_i$ .

Pentru  $n = 0, 1, \dots, 9$  realizați:

- calculați  $P_{S,n}$  și vizualizați graficul funcției polinomiale asociate lui  $P_{S,n}$ .
- calculați și vizualizați evoluția erorii  $E_{S,n}$  pe mulțimea de antrenare în care predicțiile se realizează folosind  $P_{S,n}$ .
- calculați și vizualizați evoluția erorii  $E_{T,n}$  pe mulțimea test (încărcați exemplele test din fișierul *exempleTest.mat*) în care predicțiile se realizează folosind  $P_{S,n}$  (folosiți funcția Matlab *polyval.m*).
- împărțiți mulțimea de antrenare inițială  $S$  în două mulțimi  $S_1$  (noua mulțime de antrenare – 7 exemple) și  $S_2$  (mulțimea de validare – 3 exemple). Antrenați modelul  $P_{S_1,n}$  pe  $S_1$  și testați performanța modelului pe mulțimea  $S_2$ . Alegeți din cele 10 de funcții polinomiale posibile pe cea cu cea mai mică eroare pe mulțimea de validare.