TALLER DE PROCESAMIENTO DE SEÑALES

2do Cuatrimestre 2026 - Trabajo Práctico Nº 2 - Regresión Polinómica

---

Se desea predecir el consumo de combustible de un auto (en millas por galón) a partir del peso y la potencia del auto.

(**a**) *Creación de la base de datos:*

1. Descargar la base de datos `https://archive.ics.uci.edu/dataset/9/auto+mpg`.
2. Construir la base de datos a partir de las columnas:
   - `horsepower`: potencia del motor en caballos de fuerza (feature 1).
   - `weight`: peso en libras (feature 2).
   - `mpg`: eficiencia en millas por galón (target).
3. Descartar las muestras con valores faltantes.

(**b**) *Sin Regularización*

1. Crear una clase `PolinomialRegresion` con los metodos `fit(self,x,y)`, `predict(self,x)`, `predict_rmse(self,x,y)` (raíz del error cuadrático medio) que aplique regresion polinomial, con el grado como argumento en el constructor. Se sugiere usar internamente las clases `Pipeline, LinearRegression, PolynomialFeatures` (sklearn).
2. Agregar un método `plot` que grafique las curvas de nivel de predictor, junto con las muestras de entrenamiento. Es decir, a partir de una grilla regular, hacer un gráfico utilizando `contourf` (matplotlib) donde los ejes sean los predictores (features) y las curvas de nivel coloreadas correspondan a la estimación.
3. Utilizando `train_test_split` (sklearn), separar los datos en dos conjuntos: entrenamiento y testeo con proporciones 75 % y 25 %. Evaluar el RMSE del predictor lineal, para training y test. Comparar con el desvio estandard del target (explicar por qué este esta es una referencia razonable). Graficar.
4. Evaluar el RSME de training y test para polinómios de grados 1 a 7. Graficar para grado 3.

(**c**) *Con Regularización*

1. Definir una nueva partición del conjunto de datos con las proporciones $[70\,\%, 15\,\%, 15\,\%]$ (train, validacion, test).
2. Crear una clase `PolinomialRegresionRidge` que modifique `PolinomialRegresion` utilizando `Ridge` (regresión lineal con regularización L2) de sklearn.
3. Sea α el hiperparámetro que controla la regularización. Computar el RSME sobre el set de validación, para diferentes valores de $\alpha = a \cdot 2^b$, (valores sugeridos $a = 0.03$ y $b = 0, 1, \cdots, 10$).
4. Para el α que minimiza el RMSE de validación, evaluar el RSME sobre el conjunto de testeo.