Osman Şah Yılmaz, 31316

CS412-Machine Learning Course

Homework-2

Notebook:
https://github.com/osmansahyilmaz/CS412-HW2/blob/main/CS412-HW2-OsmanSahYilmaz.ipynb

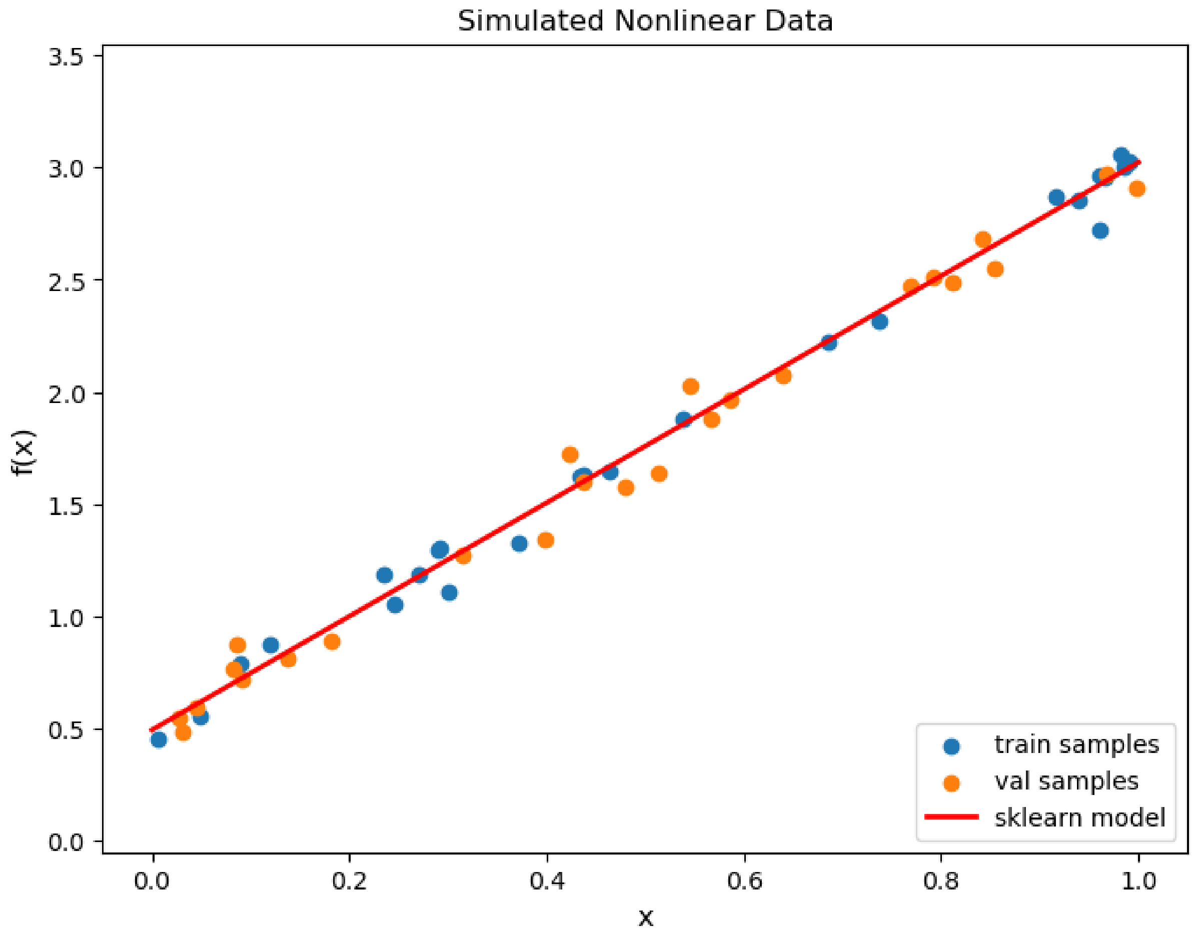
Part 1.a Results

Method: sklearn Linear Regression
Regression Coefficients:

- Intercept: 2.52838262
- Slope: 0.49447669

Mean Squared Error (MSE): 0.00795462682779033

Plot:



Part 1.b Results

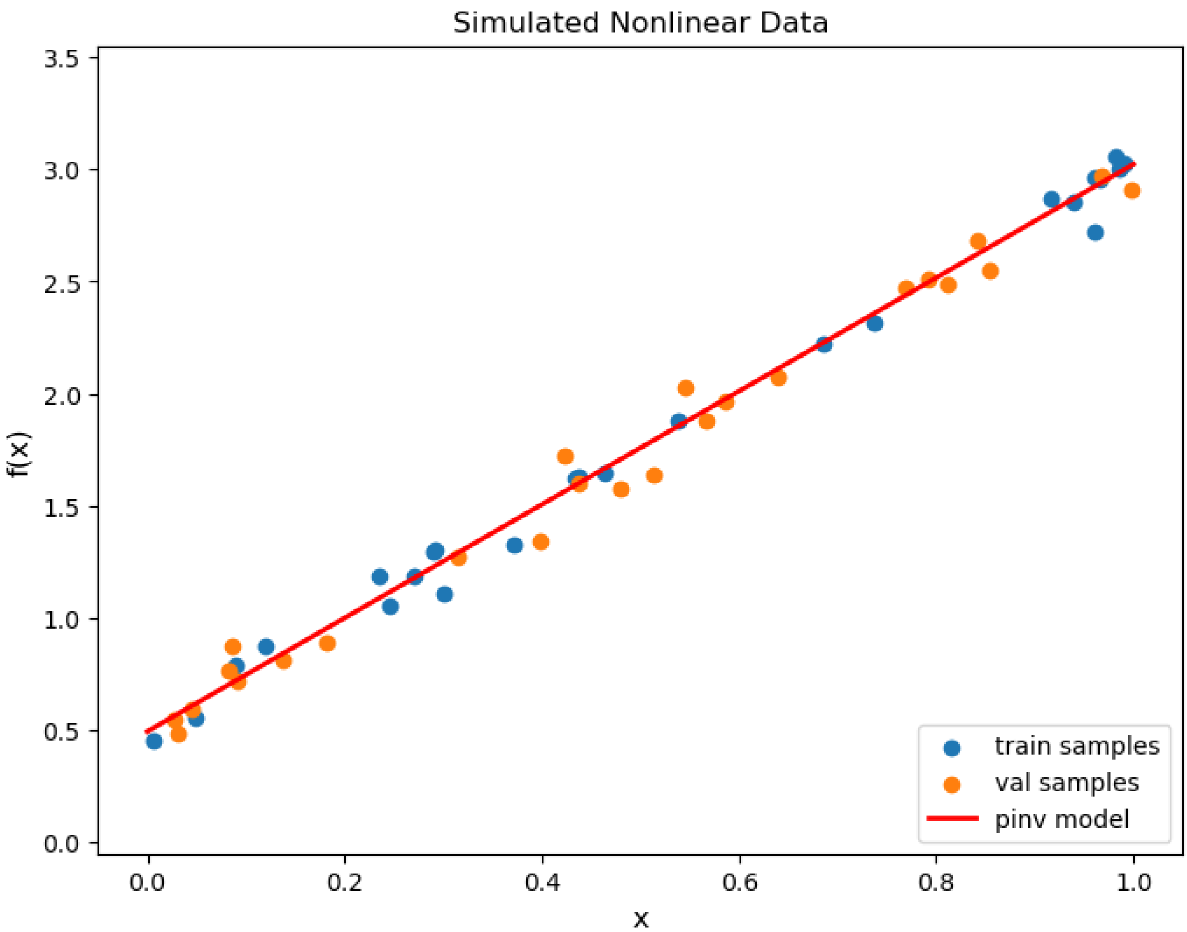
Method: Manual Solution Using Pseudoinverse
Data Matrix Shape:

- X_train_extended shape: (25, 2)
- X_val_extended shape: (25, 2)

Regression Coefficients:
(Include the coefficients computed using the pseudoinverse method)

Mean Squared Error (MSE): 0.00795462682779035

Plot:



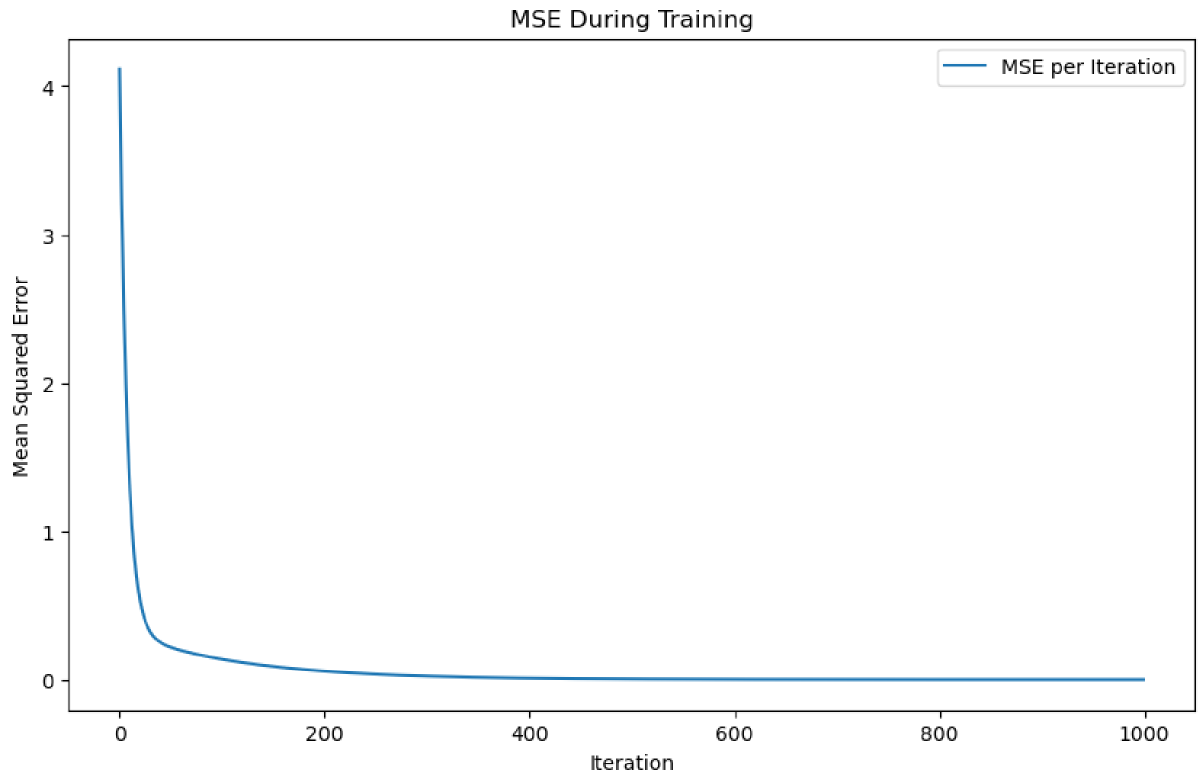
Part 1.c Results

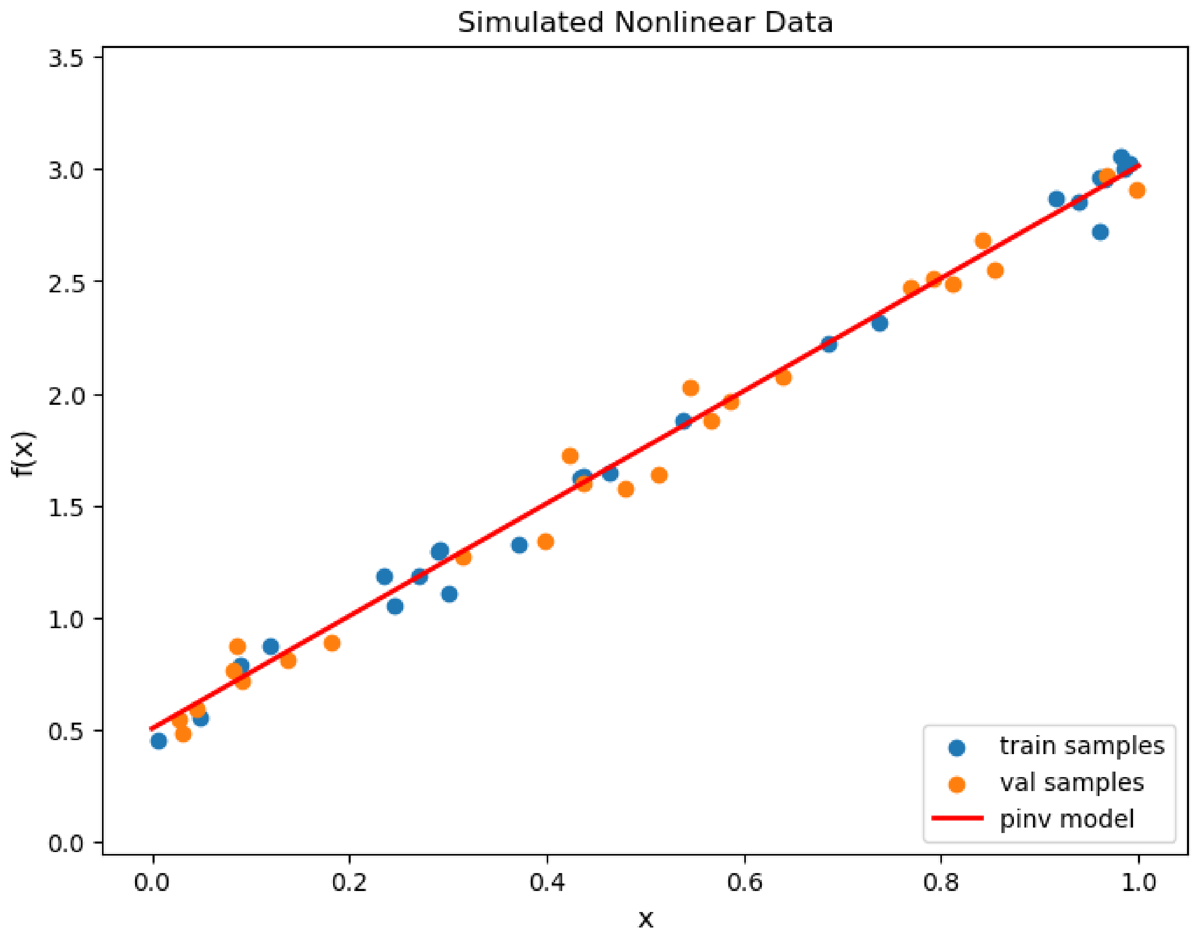
Method: Gradient Descent

MSE at Key Steps:

- Step 1: 4.1147
- Step 100: 0.1432
- Step 200: 0.0622
- Step 300: 0.0288
- Step 400: 0.0150
- Step 500: 0.0093
- Step 600: 0.0069
- Step 700: 0.0060
- Step 800: 0.0056
- Step 900: 0.0054
- Step 1000: 0.0053

Plots:





Comment:

The gradient descent solution converges to a set of coefficients that yield an MSE of 0.0053, which is very close to the MSE values obtained using both the sklearn and pseudoinverse implementations. Minor discrepancies can be attributed to the iterative nature of gradient descent (depending on the chosen learning rate and stopping criteria).

Part 2.a Results

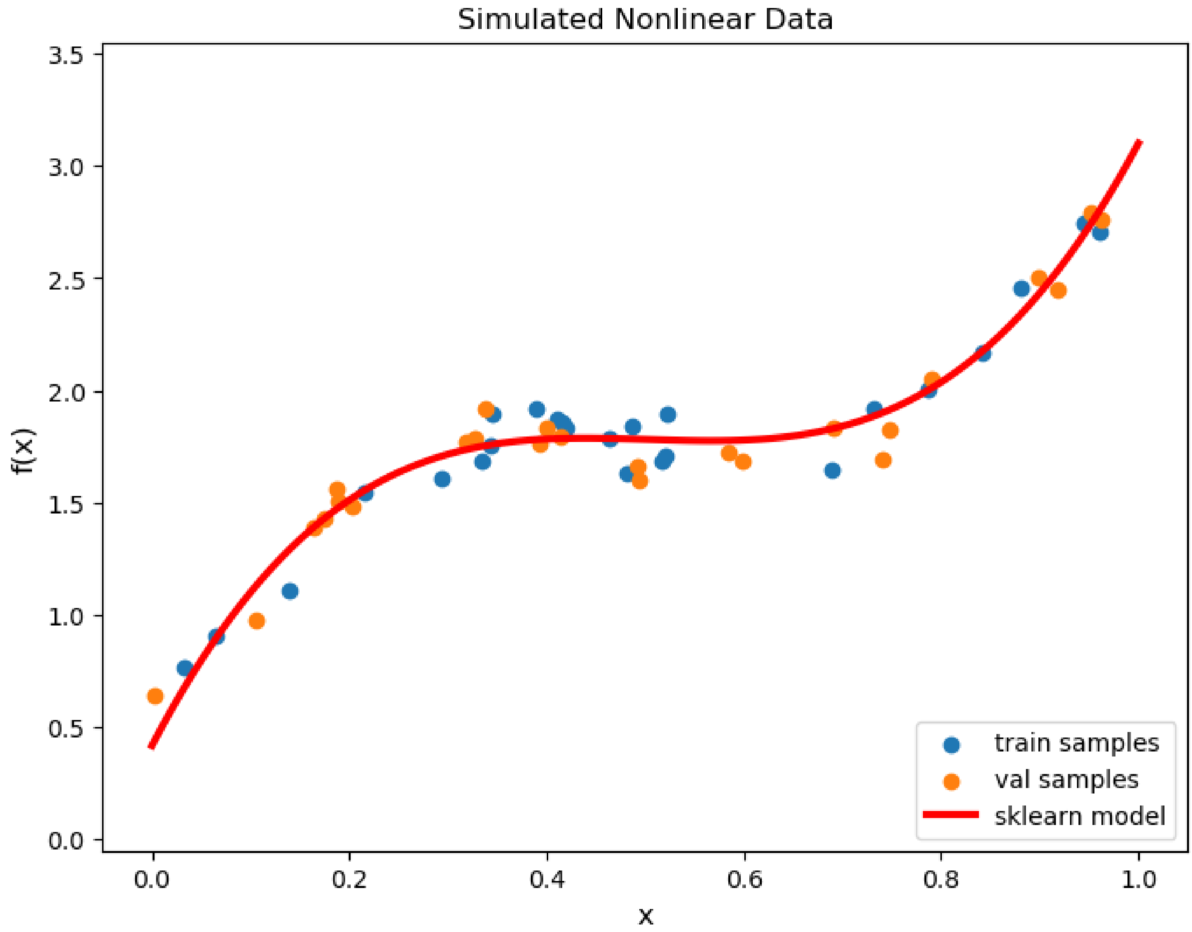
Method: Sklearn Polynomial Regression

Results for Different Degrees:

- Degree 1: MSE = 0.05830097537046252
- Degree 3: MSE = 0.009619587973839243
- Degree 5: MSE = 0.010278583809367546
- Degree 7: MSE = 0.010116792111832697

Best Model:
Degree 3 is determined as the optimal model based on the lowest validation MSE.

Plot:



Part 2.b Results

Method: Manual Polynomial Regression Using the Pseudoinverse

Approach:

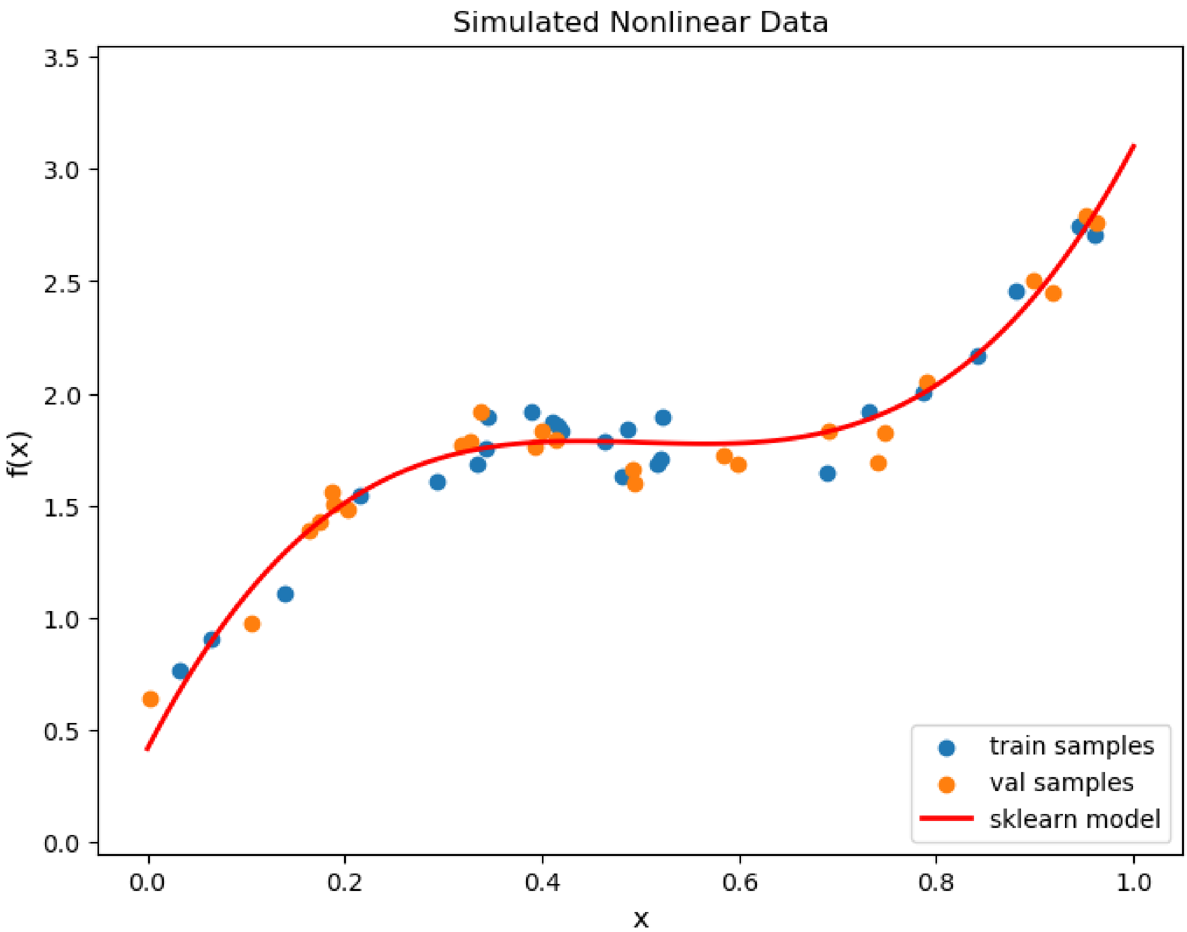
1. Construct the expanded data matrix X (including a column of ones) for the selected polynomial degree (degree 3 in this case).
2. Compute the regression coefficients using the pseudoinverse, with the equation: w = pinv(X) · y.

Data Matrix Shapes:

- X_train_manual shape: (25, 4)
- X_val_manual shape: (25, 4)

Mean Squared Error (MSE): 0.009619587973839395

Plot:



General Discussion on the Degree Parameter (Part 2)

Effect of the Degree Parameter:

• Degree Too Small (e.g., Degree 1):

- The model is too simple and fails to capture the nonlinearity present in the data, leading to underfitting and a higher MSE.

• Degree Too High (e.g., Degrees 5 and 7):

- Although a higher degree increases model flexibility and reduces training error, it can lead to overfitting by capturing noise rather than the underlying trend. In these cases, the validation MSE does not decrease further and may even worsen.

• Optimal Degree (Degree 3):

- A polynomial degree of 3 strikes a balance between bias and variance. With a low MSE (approximately 0.00962), it captures the necessary nonlinearity without overfitting, thus providing better generalization to unseen data.

Comparison of Methods in Part 2:

Both the sklearn implementation using PolynomialFeatures with LinearRegression and the manual pseudoinverse solution achieve nearly identical performance (MSE ≈ 0.00962) for degree 3. This consistency confirms that both approaches are effective when using the optimal polynomial degree. The key finding is that selecting the appropriate degree parameter is critical to balance underfitting and overfitting.