Assignment on Weighted and Locally Weighted Linear Regression

Theoretical Questions:

1. **Matrix Form of the Cost Function:** Derive the matrix form of the cost function for Weighted Linear Regression. Explain how the weights are incorporated into the cost function.

2. **Generalized Normal Equations:** Derive the generalized normal equations for Weighted Linear Regression. Show how these equations can be used to find the optimal weight vector θ.

3. **Maximum Likelihood Estimation with Weights:** Explain how Maximum Likelihood Estimation (MLE) can be used to derive the cost function for Weighted Linear Regression, assuming a Gaussian distribution for the errors with varying variances (heteroscedasticity) corresponding to the weights.

Coding Problem: Implementing Locally Weighted Linear Regression (LWLR)

Objective: Implement Locally Weighted Linear Regression from scratch, tune its bandwidth parameter τ, and evaluate its performance on a given dataset.

Dataset: You can generate a synthetic dataset or use a publicly available regression dataset (e.g., a simple 1D regression problem with some non-linearity).

Tasks:

1. **Implement LWLR:**

* Write a function predict_lwlr(X_train, y_train, x_query, tau) that takes training data X_train, y_train, a query point x_query, and the bandwidth parameter tau as input.

* Inside this function, for each x_query:

* Calculate the weights $w^{(i)}$ for each training example $(x^{(i)}, y^{(i)})$ using the Gaussian kernel: $w^{(i)} = \exp\left(-\frac{(x^{(i)} - x_{query})^2}{2\tau^2}\right)$.

* Form the weight matrix W (a diagonal matrix with $w^{(i)}$ on the diagonal).

* Solve for the optimal parameters θ using the weighted normal equations: $\theta = (X^T W X)^{-1} X^T W y$.

* Return the prediction for x_query using the calculated θ.

2. **Tune the Bandwidth Parameter τ:**

* Experiment with different values of τ (e.g., 0.01, 0.1, 0.5, 1.0, 5.0, 10.0).

* Visualize the regression line for each τ value on your dataset.

* Discuss the effect of τ on the model's bias and variance. How does a small τ differ from a large τ?

3. **Evaluate Model Performance:**

* Split your dataset into training and testing sets.

* For each chosen τ, train the LWLR model on the training set and make predictions on the test set.

* Calculate a suitable regression metric (e.g., Mean Squared Error (MSE) or R-squared) for each τ on the test set.

* Present your results and conclude which τ performs best for your dataset and why.